DOCUMENT RESUME

ED 086 214 IR 000 063

AUTHOR	Asher, William; Vockell, Edward
TITLE	Information Quality and Educational Decision Making. Final Report.
INSTITUTION	Purdue Research Foundation, Lafayette, Ind.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-1-E-189
PUB DATE	Feb 73
GRANT	OEG-5-72-0037(509)
NOTE	158p.
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	*Decision Making; Diffusion; *Educational Research; *Information Dissemination; Information Networks; Information Services; Information Utilization; *Program Evaluation; *Quality Control
IDENTIFIERS	Educational Resources Information Center; ERIC; Research in Education; RIE

ABSTRACT

Research investigated: 1) the quality of literature disseminated through an educational information channel, 2) its acceptance level, and 3) its effects upon decision-making. Some 102 research documents were selected from the 1971 entries in Research in Education (RIE), the abstracting and microfiche dissemination service for unpublished literature compiled by the Educational Resources Information Center (ERIC). Researchers assessed the documents' quality and educational decision-makers evaluated their practical utility. Results showed that the researchers rated the overall quality of the RIE documents as low, but that their accpetance by decision-makers was high. Decision-makers with less research sophistication tended to overrate the quality of the documents; the quality of the materials did not differ among the ERIC Clearinghouses, but was found to vary directly with the quality control level of the organizations which sponsored the research. The study, therefore, indicated that RIE often disseminated low quality information, which was likely to have harmful effects upon its intended audience. A system of rapid refereeing was suggested as a means of maintaining the positive advantages of RIE while reducing the negative consequences. (Author/PB)

FILMED FROM BEST AVAILABLE COPY

ED 086214

EM

FINAL REPORT
Project No. 1-E-189
Grant No. OEG-5-72-0037(509)

INFORMATION QUALITY AND EDUCATIONAL DECISION MAKING

William Asher and Edward Vockell

Purdue Research Foundation
West Lafayette, Indiana 47907

February 1973

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE
Office of Education
Bureau of Research

IR 000 063

ED 086214

Final Report

Project No. 1-E-189

Grant No. OEG-5-72-0037(509)

Information Quality and Educational Decision Making

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

William Asher and Edward Vockell

Purdue Research Foundation

West Lafayette, Indiana 47907

February 1973

The research reported herein was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

TABLE OF CONTENTS

	Page
LIST OF TABLES	vi
LIST OF FIGURES	vii
ABSTRACT	viii
INTRODUCTION	1
Need for the Study	1
Overview of the Study	1
Significance of the Study	3
REVIEW OF THE LITERATURE	5
Dissemination of Scientific Information	5
Quality of Educational Information	6
Quality Control in Research Dissemination	9
Computerized Information Dissemination Systems	13
Educational Resources Information Center (ERIC)	14
Quality of the RIE Research	16
Research Evaluation Instruments	18
Statement of the Problem	19
PROCEDURES	21
Overview	21
The Documents in the Study	22
The Educational Decision Makers	23
The Research Specialists	25
Evaluation Instruments	26
Evaluation of Documents by Decision Makers	27
Evaluation of Documents by Research Specialists	28
Reliability of the Rating Scales	30
Analysis of the Data	31
RESULTS AND DISCUSSION	36
Returns of the Evaluation Forms	36
Reliability and Intercorrelations of the Rating Scales	36
Quality of the Documents	37
Research Specialists vs. Decision Makers	39
Range in Sophistication of Decision Makers	42
Relation of Findings to Decision Making Processes	46

Page

Variations in Degree of Quality Control of the Reports 47
The Decision Setting 48
Summary of Results and Discussion 49

CONCLUSIONS AND RECOMMENDATIONS 51

Effects of Rapid Dissemination 51
Rapid Dissemination with Quality Control 53
Areas for Further Research 57
Conclusion 59

LIST OF REFERENCES 61

APPENDICES

Appendix A: The Documents in the Study 65

Appendix B: Letter for Initial Contact with Educational Decision Makers 71

Appendix C: The Decision Makers 73

Appendix D: The Initial Letter to the Research Specialists . 78

Appendix E: The Educational Research Specialists 80

Appendix F: Cover Letter Sent to Decision Makers with Documents to be Reviewed 81

Appendix G: Sample of Follow-up Telegram 83

Appendix H: Cover Letter Sent to Research Specialists with Document to be Reviewed 84

Appendix I: Information Quality and Uses Form 86

Appendix J: Form for Evaluating Quality of Research 91

Appendix K: Supplementary Evaluation Checklists Sent to Research Specialists 96

Appendix L: General Information Form 104

Appendix M: Tables and Figures 105

LIST OF TABLES

Table		Page
1	Quality Ratings of RIE Documents (Asher & Vockell, 1970) ..	17
2	Reliabilities of the Rating Scales	106
3	Intercorrelations of the Rating Scales	107
4	Comparative Ratings of Research Specialists and Decision Makers on Rating Scales	108
5	Means and Standard Deviations on the Rating Scales	110
6	Comparison of the Research Specialists Ratings on the Accept-Revise-Reject Scale with Those in Wandt's (1967) Study	111
7	Ratings of Research Specialists Compared to Ratings of Decision Makers	112
8	Analysis of Variance by Educational Level of Decision Makers (Ph. D. vs. Non-Ph. D.)	115
9	Analysis of Variance by Institutional Affiliation of Decision Maker	120
10	Analysis of Variance by Clearinghouse	125
11	Analysis of Variance by Sponsorship of Reports	128
12	Analysis of Variance by Decision Situation	130
13	Summary of the Characteristics of the ERIC Documents	135
14	Reasons for "Revise" and "Reject" Ratings	136

LIST OF FIGURES

Figure		Page
1	Interaction of Level of Education of Decision Maker with Discrepancies in Ratings on Rating Scale A	137
2	Interaction of Level of Education of Decision Maker with Discrepancies in Ratings on Rating Scale B	138
3	Interaction of Level of Education of Decision Maker with Discrepancies in Ratings on Rating Scale C	139
4	Interaction of Level of Education of Decision Maker with Discrepancies in Ratings on Rating Scale D	140
5	Interaction of Level of Education of Decision Maker with Discrepancies in Ratings on Rating Scale E	141
6	Interaction of Institutional Affiliation of Decision Maker with Discrepancies in Ratings on Rating Scale A	142
7	Interaction of Institutional Affiliation of Decision Maker with Discrepancies in Ratings on Rating Scale B	143
8	Interaction of Institutional Affiliation of Decision Maker with Discrepancies in Ratings on Rating Scale C	144
9	Interaction of Institutional Affiliation of Decision Maker with Discrepancies in Ratings on Rating Scale D	145
10	Interaction of Institutional Affiliation of Decision Maker with Discrepancies in Ratings on Rating Scale E	146
11	Interaction of Decision Setting with Discrepancies in Ratings on Rating Scale A	147
12	Interaction of Decision Setting with Discrepancies in Ratings on Rating Scale B	148
13	Interaction of Decision Setting with Discrepancies in Ratings on Rating Scale C	149
14	Interaction of Decision Setting with Discrepancies in Ratings on Rating Scale D	150
15	Interaction of Decision Setting with Discrepancies in Ratings on Rating Scale E	151

ABSTRACT

The quality of the literature disseminated through a major educational information channel, the acceptance levels of this literature, and its effects upon decision making and implementation in the educational systems are the major considerations of this research.

One hundred and two research-oriented documents were randomly selected from the 1971 entries in Research in Education (RIE), the abstracting and microfiche dissemination service of unpublished literature compiled by the Educational Resources Information Center (ERIC). These reports were presented to a group of Educational Research Specialists to determine their quality. The same reports were submitted to a group of Educational Decision Makers for an assessment of their quality and usefulness from the practitioner's point of view. By analyzing and comparing these assessments, an attempt was made to determine: (1) the quality of the information being disseminated; (2) the acceptance levels of the reports among Educational Decision Makers; (3) the plans Educational Decision Makers make as a result of the information being disseminated; and (4) how these implementation plans are related to the quality of the information being disseminated.

The results indicated that, while there were some high quality reports in RIE, the overall quality was rated low by the Research Specialists. On the other hand, the acceptance levels of these same reports among

Decision Makers was high. Thus a significant disparity was found between the quality of the reports and their acceptance levels among educational practitioners. Specific information relating to actual decisions these practitioners would base on these reports was not adequately obtained in this study. However, it was inferred that the rapid dissemination of low quality information was actually a disservice rather than an assistance to the Decision Makers.

In addition, the results suggested that Decision Makers with lower degrees of research sophistication were more likely to overrate the quality of the research. Although no differences in quality of research were found among the various clearinghouses, differences in quality were found to be related to the sponsorship of the paper. Papers sponsored by organizations with higher quality control were of significantly higher quality.

The results, therefore, indicate that RIE often disseminates low quality information which is likely to have harmful effects on its intended audience. The major recommendation for this study is to introduce a system of refereeing into the RIE system. A rapid refereeing system would retain the advantages of RIE, while reducing the negative side effects found in the present study.

INTRODUCTION

Need for the Study

Knowledge and information in education, as well as in science and technology in general, have proliferated rapidly in recent years. Expenditures for research specifically designed to improve educational practices have increased greatly. In addition, large scale attempts are being made to apply knowledge from other fields to educational theory and practice.

This proliferation of knowledge and information has been accompanied in the educational community by considerable concern about the adequate dissemination and implementation of reports, findings, collections, summaries, etc., which have been compiled. An increase in information accomplishes little unless it is disseminated to persons who need this information for use in their professional work. Further, the dissemination of information is useful only to the extent that the information being disseminated is accurate, of high quality, and in a form which will be of optimal use to the recipient of the information.

The present study examines the quality of the literature disseminated through a major educational information channel, the acceptance levels of the literature among its recipients, and its effects upon decision making and implementation in educational systems.

Overview of the Study

Research in Education (RIE) represents the indexing and microfiche dissemination system of the Educational Resources Information

Center (ERIC). The primary purpose of RIE is to disseminate information which is not already available in some other published sources. For example, papers presented at meetings and conventions, technical reports, final reports of various funded research projects, and other reports which are not published by their authors might be disseminated through RIE.

While rapid dissemination of educational information is desirable there is considerable evidence that much of what is disseminated is of poor quality. For example, Michael (1963), Campbell and Stanley (1963), Wandt (1965, 1967), Scriven (1967), Asher (1969), Mann (1969), Caro (1971), and Vockell and Asher (1972) have suggested that quality and reliability are often lacking, even in research published in the scientific journals. If the published research literature, which is subject to prepublication editing, is of questionable quality, it would seem that unpublished research literature, which is largely unedited, would be of even lower quality.

The present study attempts to determine: (1) the quality of the research information being disseminated in RIE, (2) the acceptance levels of this information among educational decision makers, (3) the plans educational decision makers make as a result of the information being disseminated, and (4) how possible implementation plans are related to the quality of the research.

The basic procedure employed in this study was to select a random group of RIE research documents and to submit these documents for evaluation to two groups of judges: (a) highly qualified specialists in educational research, who would determine the quality of the documents; and (b) educational practitioners who represent the audience expected

to make educational decisions based on the information disseminated in the documents (i.e., the audience to whom ERIC is directed.)

Significance of the Study

If the documents which are being disseminated by RIE are of high quality, then this dissemination service is providing a valuable service to educational decision makers. Rapid delivery of high quality research would enable educational practitioners to bring relevant information to bear on their problems much sooner than would be the case if they would have to wait for formal publication of the research. Thus practitioners could make use not only of information published by "big name" educators and authors, but also of well-evaluated practical implementations of research by other practitioners.

If the research disseminated by RIE is found to be of low quality, a typical reaction to such a finding would be to argue that such dissemination is a waste of time and money. Educational practitioners can gain little from information which is trivial or inaccurate. It will be argued in this study, however, that such dissemination of low quality information is not merely a waste of time and money, but actually an adverse influence upon intelligent educational decision making.

When a person buys a new house, he does not have to be an architectural expert: he feels he has a right to rely upon the judgment of an architect to provide appropriate, sound, architectural engineering. The situation is similar for users of educational research. A person faced with making an educational decision has the right to assume that the research information with which he is provided as a basis for his decision is accurate. He should not have to be an expert on multivariate analysis of variance to be able to make decisions about studies

which use this analytical procedure. If he is faced with a decision and has information available from a factor analytic study, his problem should not be to determine whether the factor analysis was properly done, but rather to determine whether the results of the study are applicable to his situation and what effect these results should have upon his decision.

The dangers inherent in disseminating low quality research information, therefore, are accentuated to the extent that it can be shown that educational decision makers are unable to differentiate between good and bad research. The present study attempts to examine both the quality of the research disseminated through RIE and the likelihood that low quality research which is disseminated will be considered high quality information by educational decision makers.

REVIEW OF THE LITERATURE

Dissemination of Scientific Information

A National Academy of Science Committee (1970) has listed five levels to which scientific information needs to be disseminated: (1) the scholar, (2) the practitioner, (3) the elementary or secondary school teacher, (4) the policy-maker or administrator, and (5) the citizen. Similarly, Brady and Branscomb (1972) list, in decreasing order of specialization and increasing order of breadth and generality, four general communities in need of technical information: (1) the scientific specialist (the researcher), (2) the industrial engineer (the applier), (3) the planner, the policy maker, and the manager (the innovators and guiders), and (4) the public (the consumer, beneficiary, and victim). All persons in these various categories have a right and need for such information, but they do not all need it at the same level of specification, detail, specialization, understanding, or speed. In view of such varying levels of need and sophistication, it has become increasingly obvious that some outside organizational help, beyond the usual abstracts, yearly reviews, summarizing books, etc., is necessary for most users of scientific information to help them keep abreast of new developments in various fields of science and for their various purposes.

Dissemination of research findings can be either formal or informal. Informal dissemination includes such procedures as discussions with colleagues, corresponding among interested parties, and conversations

and exchanges at meetings. Garvey, Lin, and Nelson (1970) present some interesting comparisons between such informal presentations in the physical sciences, the social sciences, and education and their transition into more formal dissemination through journals. They find that patterns of information flow somewhat differently in these various disciplines. The communications structures in education and the social sciences are relatively incohesive as compared to those of the physical sciences.

In formal dissemination of scientific information through journals, a major role of review and quality-control has traditionally been allocated to an editor or board of editors or consultants. The National Academy's Committee (1970) states: "The editor's task is to decline work which is duplicative, incompetent, incorrect, or totally pedestrian. This set of editorial judgments is the backbone of the scientific information system. It protects the inexpert reader and those who provide research funds, while assuring scientists in the field that the published work has been performed with competence and that the findings are probably reliable." (p. 413)

Quality of Educational Information

Havelock (1969) presents a detailed study of the processes of innovation, dissemination, and knowledge utilization. He also examines the characteristics of the individuals and organizations which facilitate or inhibit these three processes. Although this and similar studies reviewed by Havelock provide essential information about the systems and linkages involved in the dissemination process, there is little explicit consideration of the quality of information being disseminated; that it may be inaccurate or even have detrimental effects in the

dissemination processes as the result of such inaccuracies.

There has long been criticism of the adequacy of the research and evaluation literature published in the education journals. Speaking in general of the era of pre-USOE research support and dissemination, for example, Michael (1963), an editor of the Research Methodology issue of the Review of Educational Research, has suggested, perhaps somewhat harshly, that, "Probably, on the average, only 10 percent of published papers in educational journals are worthy of being reported in the Review." Similarly, Scriven (1967) has maintained that in education "...by minimum acceptable research standards, 95 percent of the work in the field...that is concerned with causal analysis is, either by theoretical or practical standards, invalid or trivial." Sieber (1968) reports a review of 250 comparisons of live and televised instructions contained in 31 reports which showed that most of the comparisons were ill designed, used inadequate samples, misinterpreted the data, or suffered from other serious flaws.

Wandt as Chairman of an AERA Committee on Evaluation of Research (1965, 1967) published a report on an empirical study of the quality of research in education. He selected 125 articles which he determined to be representative of the research articles published in the broad field of education journals in 1962. These also included associated fields of child development, educational psychology, and sociology. He submitted these articles to 125 judges who were deemed experts in the field of educational research by an AERA panel. The judges were asked to rate the articles "accept," "revise" or "reject" on the basis of their acceptability for publication in a journal of educational research. Of these published articles, the judges rated

19% "accept," 41% "revise," and 40% "reject." The judges also gave rather detailed lists of specific shortcomings which occurred in the articles which they felt should be rejected or returned for revision.

Campbell and Stanley (1963, p. 176) suggest, "Much research in education today conforms to a design in which a single group is studied only once, subsequent to some agent or treatment presumed to cause change....such studies have such a total absence of control as to be of almost no scientific value." Glass (1968) states that unless evidence of the worth of a new procedure generated from a single case or group is supported by public and replicable evidence, then the support for the ideas and procedures can only be appeals to authority. These authorities may be very able people, but it leaves the door open to self-interested persons, quacks, and frauds who to the public may seem to have equal, or even greater authority. Corey's book, Action Research to Improve School Practices (1953) points out, "The results of many educational experiments have been controversial because inadequate provision was made for obtaining data describing their success or failure. Those who substitute for the 'tried and true' newer and presumably more promising practices are under an especial obligation to obtain objective evidence about consequences (p. 100)."

More recently, Campbell (1969, p. 409) has indicated possible political involvements accompanying many educational innovations which could make accurate evaluations extremely difficult. "...most administrators wisely prefer to limit the evaluations to those the outcomes of which they can control, particularly insofar as published outcomes or press releases are concerned. Ambiguity, lack of truly comparable comparison bases, and lack of concrete evidence all work to increase

the administrator's control over what gets said, or at least to reduce the bite of criticism in case of actual failure. There is safety under the cloak of ignorance." Further, as Caro (1971, p. 91) indicates, "Those who actually carry out the programs to be evaluated are subordinate to those to whom the evaluators report. The issues addressed by the evaluation and the manner in which results are reported are strongly related to sponsorship. Consequently, the interests of the general public, practitioners, and recipients of services are not often fully served by evaluators."

Another difficulty arises out of the demand for quick results. Mann (1969, p. 13) states, "The better the study, the longer it takes, and consequently the less usefulness it may have. Conversely, the sloppier the procedure, the more likely it is to provide information on questions of interest even though this data will be of doubtful validity."

Quality Control in Research Dissemination

The implications of information quality for educational evaluation are proposed by Stufflebeam, Foley, Gephart, Guba, Hammond, Merriman, and Provus (1971, p. 61-63) in their evaluational model. It distinguishes among four general types of decision settings: (1) decisions to effect large changes supported by a high level of relevant information grasp, (2) decisions involving small changes and high information grasp, (3) decisions involving small changes and low information grasp, and (4) decisions involving large changes and low information grasp. (A change is "large" if society considers the variables to be altered important and if society considers the proposed change to be

important rather than trivial.) The depth and quality of information demanded from the dissemination process will vary from setting to setting. Decisions which involve large changes will demand more detailed, accurate information than those involving smaller changes. Stufflebeam, et al. state that for educational practitioners "evaluation must be more extensive when there is only little information (or when the client cannot use available information in its present form)."

Asher (1972) states that it is not possible for professional educators to read the literature in education and the related journals without a good knowledge of the theory of measurement, internal validity concepts of research design, and some statistics. This problem is more serious in education than in the physical and biological sciences. One reason may be that engineers and physicians are trained in the measurement principles and research methods of the physical or biological sciences, while most professionals in education are not so well trained in the behavioral sciences. In addition, comments, critiques, and rejoinders are a standard part of the literature in the physical sciences, whereas this is not the case in the behavioral sciences. Vockell and Asher (1972) determined the 15 most frequently cited journals in the 1969 Encyclopedia of Educational Research and sent questionnaires to the editors of these journals to determine policy with regard to critical comments. The results clearly indicate that these journals do not as a general practice contain critiques, comments, rejoinders, or critical letters to the editor.

Kronick (1969) states that the scholarly journal originally developed as an open record in which each scholar submitted his findings to his fellow scientists for their review and criticism. Such

exchange of critical comment is looked upon as an essential part of the scientific dissemination process. Garvey and Griffith (1971, p. 357) have reported on a study of psychologists who had distributed preprints of their scientific papers: "Over 60% of these authors received feedback that prompted them to modify their manuscripts. These modifications were not simply a matter of improvement in grammar and style of the manuscript but, instead, involved significant modifications such as reanalysis of data, redefinition of concepts, etc." The importance of good refereeing and quality control at the editorial level is magnified by the fact that teacher-oriented organizations seem to be initiating publication policies which will put "practical" information rather than research studies into the hands of classroom teachers. For example, Riedesel (1971) indicates that the future policy of The Arithmetic Teacher, a journal of the National Council of Teachers of Mathematics (NCTM), will be not to publish research articles. Implicit in such a policy change is the expectation that readers can implement what they find useful without bothering with the details of the research design.

Gideonse (1968) presented an output-oriented model of research and development elaborating and extending Guba and Clark's (1965) well known model. He attempts to classify the processes related to and necessary for change in education. The first stage is Development, which includes both invention and design of the innovation; the second stage, Diffusion, includes dissemination and demonstration. (Gideonse's definition of Demonstration includes such processes as "to examine and assess" the qualities of the innovation.) The final stage is Adoption, which includes trial, installation, and institutionalization. Here part of the objective is, "To...provide a basis for assessing the

quality, value, fit, and utility of the invention...i.e., to test." Thus Gideonse considers systematic evaluation of the material to be disseminated as an integral part of the Development and Dissemination process.

Asher (1972) states, "...lack of systematic research editing, critiques, and comments in the literature...hurts in a non-negative way as well. It is often stated that everybody knows the quality of the literature in education, and this is particularly true of doctoral dissertations. But there are well done studies in the literature of value to education that are ignored....Again the total volume of the literature is so large and generally of such poor quality in terms of its objectivity that often the strategy in using it is to implement those research ideas which occurred most recently, that is, what is most popular at the moment. This practice hardly builds enduring qualities in an educational system. Meanwhile communications and information of high quality tend to become obliterated in the mass of documents." Asher recommends as a partial solution an explicitly defined Elimination process added to the Gideonse (1968) model.

Garvey and Griffith (1971) view this problem from a slightly different viewpoint. They maintain that an informal information system must exist alongside the formal dimension, but they point out that this informal domain is not meant to have the same attributes as information published in the reputable journal literature. These authors maintain that scientists and authors need a forum where they can present half-formed ideas for critical commentary from their peers before re-evaluating them and presenting them to the public. Such information and ideas may turn out to be completely false, but as long as they are

contained in the informal domain, they should be relatively harmless. The potential difficulty is that such informal information may be disseminated in such a way that "...regardless of efforts to keep these reports from taking on the status of a formal communication, the receiver gets the impression that the reported findings are sound, and the discovery is original. This conclusion may be reached simply because the reports no longer have the characteristics of informal communication and come to resemble the formal journal article (p. 361)."

Computerized Information Dissemination Systems

The computer has introduced new advantages and opportunities for the storage and dissemination of scientific information. Such advantages, however, are accompanied by difficulties and responsibilities. The work of Garvey and Griffith (1963, 1964a, 1964b, 1965, 1966, 1967, 1971) has shown that concentration on retrieval techniques alone is not an adequate way to improve a communication system if the desired outcome is a net increase in the amount of relevant knowledge available to the working scientist (Clark, 1971). An excessive volume of information tends to swamp the reader (Licklider, 1966). Holt (1971, p. 331) points out that "The undoubted merits of computers in this kind of work seem to have made many of us overlook the fact that the problems we face are not primarily technological....It is becoming evident, however, that the diagnosis of our communication problem is mainly 'information input overload' and the strategy of coping with it that seems called for is reducing the amount of input by better control of quality."

With the increase in volume of information, there is greater difficulty for the inexperienced reader in determining what is good and what is bad in literature. Ziman (1969, p. 319) points out that "We must be able to rely on the basic accuracy and honesty of what we read in other people's papers, for we are always using their results in the construction of our own researches, and simply cannot find the time to repeat all their experiments, measurements, calculations, for ourselves....I cannot see how this innocence could be preserved against careerist pressures to publish, if there were no scrutiny by expert referees." The National Academy of Sciences Committee (1970) adds that even when a fully computerized information retrieval system is eventually developed "...the role of the editors and reviewers will remain unchanged: indiscriminate release of unedited reports to a computer network could well be even more disastrous than indiscriminate publishing would be today." Loevinger (1972, p. 9) adds that "Propogation of errors is far more detrimental to science than a moderate delay in the propogation of truth."

Educational Resources Information Center (ERIC)

An example of the application of computerized retrieval and dissemination systems to the field of education is the Educational Resources Information Center (ERIC). ERIC can be described as a national information system designed and supported by the National Institute of Education. Its purpose is to provide ready access to results of "exemplary programs, research and development efforts, and related information that can be used in developing more effective educational programs" (ERIC, 1970b). ERIC employs a system of specialized clearing houses, each of which is responsible for a particular educational

area. Current significant information relevant to education is monitored, acquired, evaluated, abstracted, indexed, and listed in ERIC reference sources. Through these reference sources, "any educator, anywhere in the country, has easy access to reports of innovative programs, conference proceedings, bibliographies, outstanding professional papers, curriculum-related materials, and reports of the most significant efforts in educational research and development, regardless of where they were first reported" (ERIC, 1970b).

ERIC has two major indexing and abstracting publications: Current Index to Journals in Education (CIJE), which indexes published information in education; and Research in Education (RIE), which indexes information which has not been formally published in a scientific or professional journal. A recent (1970a) ERIC newsletter states: "...ERIC's two principal publications, RIE and CIJE are remarkably efficient guides to what is good in literature. Putting ourselves in the shoes of potential users, we sagaciously clear only the best for inclusion in RIE. Computerized indexing of what is available in ERIC will facilitate your 'ferreting out' as no other information system has ever done."

It would seem that the RIE aspects of ERIC may present a contradiction. If the published literature is of the quality Wandt, Campbell, Campbell and Stanley, Glass, and Corey suggest that it is, what value can there be in the massive dissemination of unpublished literature, which has only minimal review by scientists and professionals?

Quality of the RIE Research

An investigation was initiated at Purdue University (Asher & Vockell, 1970, unpublished) to attempt to evaluate the quality of the information available through RIE. Twelve judges were selected who had taken advanced courses in research design, statistics, and measurement techniques. A total of 745 citations were randomly selected from RIE and were assigned to the judges. The judges were instructed to (a) determine the number of reports cited involving any evidence of empirical data to support the information included in the report, and (b) evaluate the quality of the articles designated as "research articles" according to predetermined criteria.

Of the 745 reports investigated, a total of only 155 were classified as data-oriented research. Of these, 22 were not available in either microfiche or hardback from ERIC's reproduction service and therefore could not be investigated further. Thus a total of 133 (17.3%) of the original 745 citations referred to available research reports.

The judges then were asked to undertake three forms of evaluation: (a) rank the reports on a 5-point scale according to the quality, (b) rank the reports "accept," "reject" or "revise" according to whether or not the articles should be accepted for journal publication, and (c) list the possible errors which appeared in reports listed either as "reject" or "revise." A summary of the results is given in Table 1.

As the results indicate, a total of only 6.7% of the original 745 citations (39% of the research documents) referred to reports which were supported by research which is either good or excellent. A total of only 2.6% of the original 745 citations (15% of the research reports) referred to reports which were considered worthy of publication as they

were written. These low ratings would seem to be the result of the previously cited looseness in the informal communication system and in evaluations sponsored by the same organization which is being evaluated. As suggested by Garvey and Griffith (1971), a serious problem can arise when such information is transformed into a formal communication system.

TABLE 1

QUALITY RATINGS OF RIE DOCUMENTS (Asher & Vockell, 1970)

	No.	Cum % Research Reports	Cum % All Reports
Excellent--A model of good practice	19	9.8	1.69
Good--A few minor defects	39	39.3	6.76
Mediocre--Not good, not bad	33	64.3	11.05
Poor--Some serious defects	37	92.4	15.86
Completely incompetent--"horrible"	10	100.0	17.16
	132*		
	Accept, Revise, Reject		
Accept	20	15.0	2.60
Revise	48	51.1	8.84
Reject	65	100.0	17.29
	133		

* = 1 report not rated

Dissemination is an integral and important aspect of scientific and professional enterprises in general and education and educational research specifically. ERIC is a major effort in educational research dissemination

process. Yet, the dissemination process is based in great part on the assumption that what is being disseminated is factual and worth disseminating. In the light of the research reviewed and conducted it seems that this might be a somewhat tenuous assumption.

Research Evaluation Instruments

Research has been reported on the development of various objective instruments for judging the quality of educational research and evaluation. Bartos (1969) reviews 32 such instruments. Wandt (1967) in his previously cited AERA work developed a list of 25 characteristics with a five-point quality-rating scale for each characteristic. Wandt also included an "accept-revise-reject" scale and a system for classifying individual errors pointed out by the judges. Persell (1966) used a similar combination of individual ratings of specific characteristics combined with an overall rating of the quality of the work, based on both the substance and the methodology of the research. Suydam (1968) has developed an instrument for evaluating experimental educational research reports, based on an analysis of nine general areas which results in a cumulative overall score for the report. A similar instrument was developed by Kohr and Suydam (1970) for evaluating survey research. Checklists and principles which can serve as guidelines for evaluating educational research and evaluation can be found in Farquahar and Krumboltz (1959), Borg (1963), Campbell and Stanley (1963), Mouly (1963), Rummel (1964), Travers (1964), Kerlinger (1964), Van Dalen (1966), Scriven (1967), Suchman (1967), Grobman (1968), Fox (1969), Gephart and Bartos (1969), Baxter (1970), and Stufflebeam, et al. (1971). McReynolds (1971) presents research on the reliability of such ratings of research papers.

Statement of the Problem

The central problem of this study is to determine the quality of the literature disseminated through a major educational dissemination channel, its acceptance levels among Educational Decision Makers, and the possible effects of its implementation in educational systems. Specifically, the following questions will be investigated:

1. What is the quality of the research information in this major educational dissemination channel?
2. How does this quality compare with that of refereed and edited research literature which is formally published in professional journals?
3. Does the quality of the literature vary among different ERIC clearinghouses?
4. Are RIE documents which have been submitted to some degree of refereeing of better quality than those which have been submitted to little or no refereeing?
5. What are the acceptance levels of these disseminated reports among the audience of educational practitioners, administrators, teachers, and developers?
6. Are there differences in research sophistication among Educational Decision Makers which are related to their evaluations of the research documents?
7. What implementation plans would these educational practitioners make as the result of the information disseminated in this channel?
8. How are these implementation plans related to the previously determined quality of the information disseminated?

9. **Are there differences between ratings by Educational Research Specialists and ratings by Educational Decision Makers which are related to the nature of the decision situation?**

PROCEDURES

Overview

In the present study, 102 research-oriented documents were randomly selected from ERIC's Research in Education. These documents were each assigned to an Educational Decision Maker and to an Educational Research Specialist to be read and evaluated. The Research Specialists were expected to rate the documents according to the quality of the research information, and the Decision Makers according to its usefulness in their decision making processes.

The results of the evaluations of the two groups were tabulated and compared to determine (a) the quality of the research contained in the documents, (b) whether discrepancies existed between the ratings of the two groups, (c) the kinds of decisions Educational Decision Makers would make based on these reports, and (d) the relationship between the probable decisions and the quality of the research.

In addition, the documents were subdivided into groups according to the amount of refereeing they had received prior to dissemination, and these groups of documents were compared to determine if differences in quality existed. The documents were also subdivided according to the level of sophistication of the Decision Makers who reviewed them, and the ratings of these documents were compared to determine whether the discrepancies between ratings of Decision Makers and rating of Research Specialists were related to the Decision Maker's level of sophistication.

Each of the components mentioned above will be described in detail in subsequent sections.

The Documents in the Study

A five percent random sample was drawn from all the documents indexed and abstracted in ERIC's Research in Education (RIE) during 1971. Five two digit numbers were selected from a table of random numbers, and all documents whose accession numbers ended in these two digits were included in this initial sample. This yielded a total of 617 citations. Next the abstracts of these documents were examined, and each report was classified as either research-oriented or non-research-oriented. A report was classified as research-oriented if it contained empirical data which were used either to evaluate a program or to test an hypothesis. Otherwise it was classified as non-research-oriented. Thus experiments, quasi-experiments, scientific surveys, field studies, and scientific case studies were classified as research-oriented. Bibliographies, reviews of the literature, census or status surveys, policy statements, personal opinions, curriculum guides, administrative guidelines, and other documents which contained no supporting empirical data were classified as non-research-oriented.

This classifying was done by two independent judges. Initial classifying provided agreement on all but thirty documents. The abstracts of these thirty documents were re-examined, and the full documents were obtained when available on microfiche. On the basis of the examination of this further information the same two judges again classified the documents. This second attempt at classification resulted in agreement on all but five documents. These five documents were examined by a third judge and were classified as research-oriented or non-research-oriented

on the basis of the agreement of two out of the three raters. This process of classification resulted in 203 of the original 617 citations being classified as research-oriented.

Forty-one of these 203 research-oriented documents were listed in Research in Education as not available from the ERIC Document Reproduction Service (EDRS) in either microfiche or hardcover format. Another 21 were available in microfiche but not in hardcover format. Thus of the original 617 documents in the initial sample of documents, only 141 were research documents in the format needed for mailing to the reviewers. From these, 102 documents were randomly selected and were purchased from EDRS for subsequent analysis in the current study. (The documents were numbered from 1 through 141. Thirty-nine three-digit numbers within this range were selected from a random numbers table, and the corresponding 39 documents were omitted from the study.) The titles of these 102 documents and their RIE accession numbers are given in Appendix A.

The Educational Decision Makers

The Educational Decision Makers consisted of professional educators in elementary and secondary school systems, in state educational agencies, in colleges and universities, and in other education-related fields which are part of ERIC's audience. A person was classified as a "Decision Maker" if he was an educator at a level higher than a teacher of a single classroom and conceivably could make an educational decision based on the information in the report he would review. All Decision Makers were persons who had volunteered to take part in the study and who were told that they would receive an honorarium ranging from $15.00 to $30.00 for their services. Initial contact was made with these Decision Makers either by letter (see Appendix B) or by phone. (The phone message conveyed the same information contained in the letter.)

With a few exceptions, the Decision Makers were selected from Indiana and surrounding states and metropolitan areas. This procedure was followed for several reasons. First, such a sample was geographically convenient for researchers from Purdue University, who were able to make the necessary personal follow-ups required by the study to obtain an exceptionally high rate of return and to obtain necessary clarifications and interpretations of results. Second, a population of Educational Decision Makers was readily available from up-to-date lists from state, city, and other educational agencies. Third, it appears that Indiana and geographically close major metropolitan areas should not be considered atypical of school personnel, systems, and institutions throughout the United States. Indiana is known as the "crossroads of the nation," and there is evidence from Project TALENT that Indiana high school students are quite typical of U.S. students in aptitude and achievement (Asher & Dodson, 1972, in press).

A list of Educational Decision Makers was obtained of participants in the Model Training Project (MTP) in Educational Evaluation, a nationwide program centered at Ohio State University. The MTP trains and retrains personnel in evaluation techniques to be employed as part of the decision making and change processes. Additional names of Decision Makers were obtained from the Indiana State Department of Public Instruction and from Phi Delta Kappa. These prospective participants were contacted by letter (Appendix B) to solicit their participation. Participants listed areas of professional interest, and research reports were then matched to these areas of interest.

The documents not placed through this process were then listed and attempts were made either by letter of phone to obtain reviewers who would

be interested in the specific area covered by each document. Lists from various departments of education, faculty lists from universities, and recommendations from personnel in these two groups were consulted. The potential reviewers were contacted and their participation was solicited. This procedure resulted in the eventual placement of all the documents.

The names and professional positions of all the Educational Decision Makers are listed in Appendix C.

The Research Specialists

The list of 166 reviewers employed by Wandt (1967, p. 32) was obtained, and an attempt was made to mail a letter to all of these reviewers to solicit their participation as reviewers of ERIC research-oriented documents. A copy of the letter is included in Appendix D. These reviewers had been selected by Wandt on the basis of professional expertise. At the time of his study all were AERA members, 98% had doctorates, 63% had taught research methods courses, 75% had supervised Ph.D. work, and their median number of professional publications was 16.

Thirty-five of the original 166 reviewers contacted replied that they would be unable to participate. Reasons for non-participation were not actively requested, but when these were volunteered they ranged from overly complicated work schedules to unwillingness to "participate in OE enterprises." One hundred reviewers replied that they would be willing to read and evaluate a document. Two reviewers were unable to participate, but recommended specific colleagues they felt were qualified to take their places, and these suggestions were accepted. Three reviewers were known to have died since the Wandt study. The other 26 reviewers could not be contacted or failed to respond after five months, and no further information was obtained on them.

Attempts were made to send reviewers documents close to their field of interest. In the course of the study it became apparent that several of the documents were completely outside the areas of expertise of the reviewers available. For example, none of the reviewers felt qualified to evaluate a document on nursing education. In such cases, persons from the content area of the report who were also highly qualified in research skills were asked to participate. Their qualifications as researchers were based on their published research in the designated content area and the recommendation of a Research Specialist already in the sample. Three additional reviewers were obtained in this manner. The names of all the Research Specialists are included in Appendix E.

To further ascertain the qualifications of these reviewers as Research Specialists, the reviewers were asked at the time they were completing the evaluation form (a) whether they were currently teaching a course in research methodology or had ever taught such a course, and (b) how many research reports they had published in professional journals or presented at professional meetings. The precise wording of the question used to obtain this information is included as part of the questionnaire in Appendix J.

Since the primary function of these reviewers was to read and evaluate the reports assigned to them from the point of view of research quality, these reviewers will be referred to as "Research Specialists" in subsequent portions of this report.

Evaluation Instruments

The Educational Decision Makers completed the Information Quality and Uses Form (Appendix I). The techniques employed in analyzing the responses to this evaluation form are explained elsewhere in this report. The

rating scales, which are the primary sources of data in the present study, are labeled A, B, C, D, and E. These labels were not on the copy of the evaluation form sent to the reviewers, but are added here to facilitate readability. The Decision Makers also completed the General Information Form (Appendix L).

The Research Specialists completed the Form for Evaluating Quality of Research (Appendix J). This form differs from the Information Quality and Uses Form in two ways: (1) the references to decision making are omitted, and (2) additional leading questions were added as part of item three. The rating scale items correspond to the similarly labeled items on the Decision Makers' evaluation form. These overlapping items are the main basis for comparison of the two sets of evaluations.

Both forms contained open-ended questions regarding the educational significance and possible methodological shortcomings in the document being evaluated. The Information Quality and Uses Form also contained two seven-point rating scale items concerning the nature of the decision setting and an open-ended question on possible plans for implementation of the research results.

Information on the reliability and intercorrelations of the items on the rating scales will be found elsewhere in this report. Prior to their use in the present study, all forms were administered, with satisfactory results, to an advanced educational research class to determine the clarity and readability of the formats.

Evaluation of Documents by Decision Makers

A copy of one of the documents in the sample, matched to area of professional interest, was mailed to each of the Educational Decision Makers who had volunteered to participate. The Decision Makers

were instructed to read the document and to return the completed evaluation form within a week to ten days. A copy of the cover letter which accompanied the documents is contained in Appendix F.

If the evaluation of a document was not returned within three weeks after the document had been sent to the Decision Maker, he was contacted either by telegram or by phone. At this time he was asked if he could complete the evaluation in the immediate future. If this was impossible, he was asked to return the document so that it could be sent to a different reviewer. A sample telegram is contained in Appendix G. If a document was returned unreviewed by the Decision Maker, then the participation of a new Decision Maker interested in the content of the document was solicited according to the same procedures employed in the selection of the original reviewer, and the document was immediately sent to this new reviewer.

Evaluation forms were returned by mail in stamped, self-addressed envelopes which accompanied the document when it was sent to the reviewer.

Evaluation of the Documents by Research Specialists

Each of the Research Specialists received a document to read and evaluate. Since the lists of both the documents and the Research Specialist were predetermined, it was much more difficult to match areas of interest in the case of the Research Specialists than in the case of the Decision Makers. The Research Specialists were instructed to return the document they had received if they found it was entirely outside their area of professional interest and expertise. When documents were returned for this reason, the document was reassigned to different Research Specialists. Since the number of available Research Specialists was limited, this procedure made it necessary for several readers to read and

evaluate a second document. An attempt was made to reassign a different document to reviewers who had returned their original document, but if this was impossible it became necessary to drop the Research Specialist from the sample of reviewers. (In such cases, the names of the Research Specialists who were thus eliminated are excluded from Appendix E.)

In addition to the evaluation form contained in Appendix J, the Research Specialists received three supplementary evaluation checklists (Appendix K) which they were encouraged to consult to whatever extent they felt advisable. These supplementary checklists (which were color-coded to reduce confusion with the evaluation form to be completed as part of the study) were for the benefit of the reviewers and were not to be returned with the evaluation form. These supplementary checklists were sent only to the Research Specialists, not to the Decision Makers. It was felt that sending these to the Decision Makers would artificially focus their attention on methodological concerns, whereas what was desired was a "typical" reaction of the Decision Maker. Specific focusing on methodology was felt to be beneficial in the case of the Research Specialist, since this was their area of expertise.

A copy of the cover letter which accompanied the document is contained in Appendix H.

If a document was not returned within three weeks, the reviewer was contacted by telegram or telephone. This procedure was identical to that for the Decision Makers, with the exception that if a document was returned, an attempt was made to reassign the document to a Research Specialist already within the sample, rather than soliciting the participation of an additional reviewer.

Evaluation forms were returned by mail in stamped, self-addressed envelopes which accompanied the document when it was sent to the reviewer.

Reliability of the Rating Scales

Ten of the documents were initially selected for the reliability sample. Each of these documents was sent to a second Decision Maker and to a second Research Specialist. The instructions for these reviewers were the same as for the first set of readers. Correlation coefficients were computed between the ratings of the first group who read each document and the group who read the same documents as part of the reliability sample.

Two additional documents were inadvertantly sent to a second Research Specialist, and one additional document was sent to a second Decision Maker. (This happened because reviewers who had originally said they would be unable to return their completed evaluation forms by the designated deadline, eventually did return their forms.) These documents were included as part of the reliability sample. Thus, the reliability sample actually consisted of twelve documents evaluated by two Research Specialists and eleven documents evaluated by two Decision Makers.

The documents used in the reliability study are accompanied by asterisks in Appendix A. Similarly, the names of the Decision Makers and Research Specialists who took part in the reliability study are accompanied by asterisks in Appendices C and E.

The second documents in the reliability sample were used only in the computation of reliability coefficients. It was felt that double use of these documents would result in giving inappropriate additional weight to them. Thus, each of the 102 documents was used only once in all non-reliability analyses.

Analysis of the Data

The primary data to be analyzed were the ratings by the Decision Makers and by the Research Specialists on the Rating Scales at the ends of their evaluation forms. These data were analyzed both by frequency counts of the number of responses in each category on each scale and by analyses of these ratings after they had been converted to an appropriate 1-2-3-4-5 or 1-2-3 interval scale.

It is important to note that the direction of rating is reversed on Scale A of both the evaluation forms. A rating of "5" on this scale is an indication that the reviewer viewed the document to be of high quality. On all other scales, a higher rating indicates lower quality in the document being reviewed. Thus "1" was applied to "Accept," and "3" was applied to "Reject." To prevent confusion, footnotes indicating this irregularity accompany all figures and tables where Scale A is involved in the analysis.

The data were coded and transferred to computer cards. Analyses were performed on the CDC 6500 computer at the Purdue University Computing Center. Frequency counts were performed using the BMD 04D and FREQ programs (Dixon, 1968). The analysis of variance program was ANOVAR (Veldman, 1967). ANOVAR is suitable for both one-way and repeated measures analysis of variance, and yields as output means for groups and trials, Mean Squares, F-ratios, degrees of freedom, and probability levels. Means, standard deviations, and correlation coefficients were computed through the BMD 03D program (Dixon, 1968). The tables in Guilford (1964) were consulted to determine significance levels of correlation coefficients and *t* tests.

To answer the first research question (What is the quality of the RIE research literature?), frequency counts and means were computed for each of the rating scales. This was performed separately for the Research Specialists and for the Decision Makers, providing ratings of the quality of the research information as viewed by both groups of reviewers.

To answer the second research question (How does this quality compare with that of published literature?), the mean ratings on the Accept-Revise-Reject Scale were compared to the parallel data compiled by Wandt (1967) through a *t* test.

To answer the third research question, the reports were grouped according to the clearinghouses from which they originated. Clearinghouses with fewer than five research documents were omitted from further analysis. The scores from the remaining twelve clearinghouses were analyzed through a one-way analysis of variance.

To answer the fourth research question (Does degree of refereeing make a difference?), the reports were divided into five groups on the basis of the sponsorship of the research: (1) American Educational Research Association (AERA) papers, (2) papers from other conventions, (3) papers from the Educational Testing Service, (4) masters or doctoral theses, and (5) papers falling in none of the other categories. The scores from the resulting five groups were analyzed through analysis of variance.

To answer the fifth research question (Do Decision Makers and Research Specialists differ in their ratings of the quality of the research?), the frequency counts of the five rating scales of the Research Specialists and the Decision Makers were compared through Chi Square analysis. The scores of the two groups on these scales were also compared by means of a repeated measures analysis of variance.

Two judges examined the documents which were rated "Revise" or "Reject" by the reviewers to determine the basis for this rating. These documents were assigned to broad categories of reasons for rejection or revision on the basis of information provided in the open-ended items on the questionnaires. These categories for the Research Specialists and for the Decision Makers were then compared to determine whether differences in the reasons behind such ratings existed.

In addition, two groups of documents were selected for more detailed analysis: (1) reports rated "Accept" by the Research Specialists and "Reject" by the Decision Makers, and (2) reports rated "Reject" by the Research Specialist and "Accept" by the Decision Maker. Specific reasons for rejection were examined for these two groups of documents.

To answer the sixth research question (Are differences in research sophistication of Decision Makers related to their ratings?), the reports were divided into those which had been rated by a Ph.D. Decision Maker and those which had been rated by a non-Ph.D. Decision Maker, and the resulting groups of reports were submitted to a repeated measures analysis of variance.

In addition, the documents were divided into four groups according to institutional affiliation: (1) those evaluated by a Decision Maker who was affiliated with a major research organization or a person designated as a research specialist for an educational institution or system, (2) those evaluated by a Decision Maker affiliated with a college, university, or technical school, (3) those evaluated by a Decision Maker affiliated with a state or local education system, and (4) those evaluated by a person affiliated with a single elementary school, high school, or nursery school. These resulting groups of reports were submitted to repeated measures analysis of variance.

To answer the seventh research question, an attempt was made to divide the reports into four groups on the basis of the recommendations concerning implementation suggested by the Decision Maker: (1) high recommendation for implementation, (2) qualified recommendation for implementation, (3) recommendation not to implement, and (4) assertion that no recommendation to implement or not to implement could be made on the basis of the report. These classifications were derived from separate readings and categorizations by two judges of the question, "What recommendations would you make with regard to the implementation contained in this report?" from the Information Quality and Uses Form. This attempt at classification was eventually found to be impractical. Consequently, no direct analysis was possible. However, inferences were drawn from other information provided by the Decision Makers in an attempt to answer this research question. The reasons for rejecting the attempt at classification and the basis of the inferential approach are explained in the Results section of this report.

For similar reasons, a direct approach to the eighth research question was impossible. The procedure of drawing inferences from other information provided by the Decision Makers was again followed in this analysis.

To answer the ninth research question, the reports were divided into four groups according to the importance of the decision and the degree of relevant knowledge of the Decision Maker who evaluated the document: (1) important decision with a high degree of knowledge, (2) important decision with a low degree of knowledge, (3) minor decision with a high degree of knowledge, and (4) minor decision with a low degree of knowledge. This information was derived from questions 6 and 7 of the Information Quality and Uses Form. Any response falling at four or below on the 7-point

scale was classified as "unimportant decision" or "low knowledge." The resulting four groups were analyzed through repeated measures analysis of variance.

Information relating to years of experience of the Decision Makers and number of publications of Research Specialists was tabulated, and means and medians were computed.

RESULTS AND DISCUSSION

Returns of the Evaluation Forms

Completed evaluation forms were obtained for 103 of the 113 documents (including the reliability documents) sent to Decision Makers. Completed forms were obtained for 111 of the 114 documents sent to Research Specialists. Thus the percentages of return were 91.1% for the Decision Makers and 97.3% for the Research Specialists. In cases where responses were totally ambiguous, or in the few cases where the wrong evaluation form was returned, telephone follow-ups were made to obtain the desired data. In cases where isolated items were left blank, these were treated as "No response" in the analyses.

Ratings of the same document by both a Decision Maker and a Research Specialist were obtained for 90 of the 102 non-reliability documents.

Reliability and Intercorrelations of the Rating Scales

The inter-reader reliabilities of the scales (based on the subsample of twelve documents) were low, as might be expected of 5-point and 3-point scales. None were significant at the .05 level. The lowest reliabilities were those for the Accept-Revise-Reject scales for the Research Specialists sample (.15) and for the Educational Decision Makers sample (.24). The other reliabilities ranged from .27 to .56. A complete listing of the reliabilities is included in Table 2 of Appendix M.

Such low inter-reader reliabilities in ratings of research reports is not rare. Editors of journals are often faced with high rates of disagreement among their reviewers, and Wandt(1967) reported similar low

reliabilities. Nunnally (1967) points out that while extremely high reliabilities are often essential in applied settings, these would often be wasteful in basic research. In the present study, higher reliabilities could probably have been obtained by using a lengthier and more detailed scale, such as the complete scale designed by Persell (1966). However, in view of the limited budget and the limited amount of time reviewers were expected to spend in evaluating the documents, the present short checklists were preferred in spite of their low reliabilities. Since the rating scales were to be used to compare groups of documents rather than individual documents, the obtained reliabilities were deemed adequate to determine significant differences and relationships.

The intercorrelations of the various rating scales (based on the full sample of 102 documents) are shown in Table 3 of Appendix M. For the Research Specialists sample, these intercorrelations ranged from .42 to .76, and for the Decision Makers from .55 to .81. All intercorrelations in both samples were significant at the .01 level. The intercorrelations of the Research Specialists' ratings with the Decision Makers' ratings on the same scales were near zero, ranging from -.13 to .19, all non-significant at the .05 level.

Quality of the Documents

The results of the rating of the documents by the Research Specialists are shown in Appendix M, Tables 4 and 5. Forty-four percent of the reports were rated as either Poor or Completely Incompetent in the overall rating (Scale A). The Research Specialists recommended that 8% be accepted for publication, that 38% be revised, and that the other 54% be rejected. (Reasons given by the Research Specialists for rating a document Revise

or Reject are discussed in the succeeding section of this report, where these reasons are compared to those provided by Decision Makers for similar ratings.) Regarding the theoretical contribution of the document to the discipline to which it pertained (Scale C), 16% were rated as either Above Average or Excellent, whereas 56% were rated as Below Average or Incompetent. The practical contribution of the document to educational practice (Scale D) was rated as Above Average or Excellent in 24% of the cases, and in 41% of the cases as either Below Average or Incompetent. Use of research methodology (Scale E), was rated as Above Average or Excellent in 24% of the reports, but as Below Average or Incompetent in 51% of the reports.

The quality of these documents can be compared on the basis of the Accept-Revise-Reject (ARR) category to the quality of the documents in Wandt's (1967) study of published educational research. Such a comparison (Table 6, Appendix M) indicates that the documents in the present study are rated as poorer in this category at the .05 level of significance. Several factors should be considered in evaluating such a comparison. First, the reviewers of the documents in the present study were drawn as far as possible from the same set of reviewers used in the Wandt study. Thus there is a basis for comparison. Second, as will be seen later in the present study, the ARR category was one of the most difficult areas in which to demonstrate significant differences. This was largely because of the small variance of the three point scale. Third, Wandt's study was based on articles published in 1962, whereas the documents in the present study were written between 1969 and 1971. This weakens the comparability somewhat. If it can be shown that the quality of published educational research literature has improved since 1962, the present analysis may

reflect a conservative estimate of the differences in quality between current published research and the research in RIE.

These ratings are supported by a high degree of authority. Seventy-six percent of the Research Specialists had taught courses in research methodology. The tabulation of their number of professional publications indicated a mean of 39.9 publications and a median of 30. It should be pointed out that this estimate of the number of publications is conservative: authors who listed "at least 30" were recorded as having 30 publications, and anyone listing over a hundred publications (Some had over 150) was recorded as having 99 publications for simplicity in computer coding. Thus any allegation that the Research Specialists were naive or out of touch with current educational trends would be hard to support.

Research Specialists vs. Decision Makers

Tables 4 and 7, Appendix M , compare the ratings of the Educational Decision Makers to those of the Educational Research Specialists. On all scales, the Decision Makers showed significant tendencies to rate the reports more favorably than did the Research Specialists. For example, the Research Specialists rated only 5 reports as "Excellent--A model of good practice," whereas the Decision Makers rated 15 reports in this category. Similarly, while the Research Specialists rated 10 documents as "Completely Incompetent," the Decision Makers rated only 1 document in this category. Likewise, the Research Specialists rated only 8 reports "Accept" and 52 reports "Reject," whereas the Decision Makers rated 30 "Accept" and only 23 "Reject." The same trend was evident on all rating scales.

Table 14, Appendix M, compares the reasons given for "Revise" and "Reject" ratings by the Research Specialists and Decision Makers. In general, the Research Specialists rejected documents on the basis of poor design or educational triviality, whereas they recommended revision when the design was unclear, when the report was too long, when there was question about the representativeness of the sample, and when the analysis appeared to be incomplete. Thus there appeared to be a clear distinction between the reasons for revision and rejection: "Revise" was recommended when shortcomings appeared which could be corrected by rewriting the report or parts of it, and "Reject" was recommended when the entire study would have to be redone. It is noteworthy that these reasons for revision indicate the same benefits which psychologists have reported as a result of submitting rough drafts of their research papers to peers prior to formal publication (Garvey & Griffith, 1971).

The dichotomy between reasons for revision and rejection was not as clear for the Decision Makers. Poor design was given with about equal frequency as a reason both for rejection and for revision. Extreme complexity or technicality, a reason for revision which did not appear among the Research Specialists, was frequently cited by the Decision Makers. In four cases the Decision Makers recommended that the report be revised, even though they made no adverse comments about it. The recommendation that the study be replicated with the reviewer's "pet theory" included as part of the study was unique to the Decision Makers. It would seem, then, that the reasons given by the Decision Makers for revision and rejection were somewhat different from those of the Research Specialists. When they recommended "Revise," this apparently sometimes meant that the entire study, or major parts of it, should be repeated with different procedures.

This meaning of "Revise" was not apparent among the Research Specialists. In addition, many of the Decision Makers' reasons for revision focused on making the report more meaningful and useful in practical situations.

Further analysis indicated that there was only one document rated "Reject" by the Decision Maker but "Accept" by the Research Specialist. This document was labeled "extremely esoteric" and "excessively technical" by the Decision Maker (who was a university professor). It is noteworthy that the Research Specialist indicated that it should be rated "Accept" from the standpoint of the state of art in the discipline to which it applied, but by best educational research standards it should be rated Revise or possibly Reject.

On the other hand, eleven documents received an "Accept" rating from the Decision Maker and a "Reject" from the Research Specialist. An examination of the reasons given for rejecting these documents indicated that one of these documents was rejected primarily because it was too lengthy and contained relatively minor problems in sampling techniques. Five documents were rejected because of serious errors in research design, and three because of the extreme triviality of the research. Two others were rejected because of a combination of poor design and triviality of the research.

One Research Specialist indicated that although the report attempted to measure change, no measure was taken of the dependent variable prior to the introduction of the experimental variable, and no control group was employed. Another Research Specialist indicated that "...what they report is in the nature of a series of mediational checks to determine whether an independent variable has been manipulated. But the question of whether that independent variable has effects upon teaching performances is not

even addressed." Another stated, "The research adds nothing to our knowledge. It is only busy work that should never have been funded. The report would not pass as a Masters thesis. Indeed, it would hardly pass as a term paper in view of the failure of the author to show a familiarity with the important work that has been undertaken in related areas." Such analysis and anecdotal information suggests that these papers which were rated "Accept" by the Decision Makers were rejected by the Research Specialists for more than minor flaws. In eleven out of the twelve cases, the concern was with extremely poor methodology or with the extreme triviality of the research, rather than with problems which would be considered a matter of taste.

These analyses indicate not only that the research reports in this study are of low quality, but also that Educational Decision Makers in general do not recognize this inferior quality. The Decision Makers showed consistent tendencies to overrate the quality of the reports they reviewed, and in many cases they rated as high quality reports with extremely serious inaccuracies. If this is true, then it would seem that ERIC's Research in Education has fallen somewhat short in its crusade to "sagaciously clear only the best for inclusion in RIE" (ERIC, 1970a). Quite the contrary, the evidence indicates that rather than providing high quality information, RIE often disseminates vast quantities of low quality information, which is very likely to be regarded as high quality by Decision Makers.

Range in Sophistication of Decision Makers

An examination of the positions of the Educational Decision Makers (Appendix C) indicates that a wide diversity exists in the qualifications

of the members of this group. It is evident that a number of these Decision Makers could legitimately be classified as Research Specialists. Although the intent of the present study was to submit the documents to a group of "typical" Educational Decision Makers, it is evident that this is not what actually took place. Since the Decision Makers in this study were volunteers, it is possible that only persons with a certain degree of confidence in their ability to read and evaluate research reports would have responded affirmatively to the introductory letter (Appendix B). In addition, many of the RIE documents dealt with higher education or teacher education, and in such cases the most appropriate Decision Makers were college or university personnel. Such Decision Makers are more likely to have some research orientation. These two biases also seemed to interact: college level personnel with research interests were more likely to volunteer to take part in a study of research quality.

Thus it seems possible that many of the Decision Makers were among the elite of this group. When an educational decision is to be made in a "typical" setting, it hardly seems likely that it will always be referred to a Decision Maker as qualified as these.

The bias of this Decision Maker sample is toward the more critical end of the scale; for example, more tendency to reject, etc. Thus the differences found in this sample would probably be markedly greater among more typical Decision Makers.

Table 8 and Figures 1 through 5 in Appendix M reflect an attempt to explore this line of reasoning further by analyzing the ratings of the documents by Decision Makers with different degrees of research expertise. It was hypothesized that holders of doctoral degrees would have a higher level of research sophistication than Decision Makers without doctorates.

The documents were divided into those reviewed by a Decision Maker with a doctorate and those reviewed by a non-doctoral Decision Maker. The ratings of the Research Specialists and the Decision Makers were then compared. If the above hypothesis were true, a significant group by trial interaction should appear. Table 8 indicates that a significant interaction was in the expected direction on the rating of practical usefulness (variable 4). Nonsignificant trends appeared on three of the other four variables (variables 1, 3, and 5). Thus, Table 8 provides some evidence that less sophisticated Decision Makers are more likely than sophisticated Decision Makers to over-rate the quality of a research report.

There are perhaps several factors which prevented the emergence of more significant results. First, it is probably naive to assume that a doctoral degree automatically implies a high degree of research sophistication. Second, a number of the Decision Makers listed as non-doctoral reviewers were at the time persuing doctorates on a part-time basis. In a few cases, the Decision Maker was a full-time Ph.D. candidate working with an educational system. Such non-doctoral Decision Makers would probably be in even closer touch with research methods and trends than would many of the doctoral Decision Makers. It seems possible, therefore, that the analysis summarized in Table 8 is not based upon as clear-cut a dichotomy of research sophistication as would be desirable. (In both this and the subsequent analysis, moreover, the low reliabilities of the scales would operate against finding significant differences.)

Table 9 and Figures 6 through 10, Appendix M, reflect a second attempt to analyze the effects of research sophistication on the ability to make an accurate determination of the quality of research. It was hypothesized

that a continuum of levels of research sophistication exists among the following four groups of Decision Makers: (1) members of research and evaluation organizations, (2) members of college faculties, (3) state and local school system personnel, and (4) faculty members of individual schools within such systems. This breakdown revealed significant interactions in the expected direction in two cases (variables 3 and 4) and non-significant trends in the other three cases.

These analyses provide evidence that Decision Makers with stronger research orientations are more likely to evaluate research in a manner which will correspond to the evaluations made by Research Specialists. However, the research dissemination needs of these research and evaluation personnel and of the college faculties are already well served by journals and other means of standard research and professional communication. On the other hand, the needs of the Decision Makers at the state, local, and individual school levels (the non-sophisticated end of the continuum) are not as adequately served by such standard professional dissemination procedures. These levels of Decision Makers predominate in our educational systems and are presumably a prime target audience for ERIC's RIE services. It is these Decision Makers who are likely to evaluate research reports most divergently from the evaluations of Research Specialists.

It may seem that these analyses belabor the obvious. It is not revolutionary to discover that persons who are more sophisticated in research make more sophisticated evaluations of research. However, this analysis does have important implications. Stufflebeam (1971, p. 63) has demonstrated that evaluation must be most extensive and most accurate when there is little information available to the practitioner who must make a decision or when the decision maker cannot use available information in its present form. If this is true, then it is not the personnel

at the sophisticated end of the research sophistication continuum who most need information by which they can evaluate accurately, but rather the non-sophisticated practitioners who have need of a great deal of information which they must accept or presume to be accurate. A useful dissemination system would be one which presents such decision-making practitioners with information which they recognize as pertinent and accurate and which is indeed of high factual quality. A dissemination system which cannot provide such a guarantee of quality to the Decision Makers at the lower end of this continuum would apparently be a disservice rather than a useful assistance.

Relation of Findings to Decision Making Processes

Originally it was intended to group the responses of the Decision Makers into varying degrees of willingness to implement the research findings of their report according to their responses to question eight on their evaluation form (What recommendation would you make with regard to the implementation of the information contained in this report?). In many cases the Decision Makers provided responses which could be appropriately classified. In other cases, however, the Decision Makers responded by specifying the kinds of further research they would like to see or revisions which could be made in the documents. Sometimes "no implementation recommended" obviously meant that the information contained in the document was useless; at other times it meant that the information was merely inapplicable to the Decision Maker's current situation; and at other times it was difficult to discern what was meant. Because of such diversity in approaches by the Decision Makers to this question and because of the ambiguity of many of the responses, the planned analysis based on this question was not implemented. Consequently, the results

analyzed in this report are confined to the perceptions of the Decision Makers towards the quality of the documents, rather than with hypothetical decisions to implement or not to implement the results of a given report.

The present study, therefore, examines the actual decision making process only to the extent that it can legitimately be assumed that a Decision Maker, given the opportunity and need, is likely to implement research which he views as high quality and to decline to implement research which he views to be low quality. This assumption appears to be valid.

Variations in Degree of Quality Control of the Reports

Although it would be unwise to generalize on the basis of the very small samples from the various clearinghouses in the present study, Table 10 of Appendix M suggests that differences in quality of documents are not related to the ERIC clearinghouses from which they originated. Thus no differences in quality control among the clearinghouses can be hypothesized from the results of this study.

Within RIE itself, however, there may be some general differences in the degree of quality control which a document might receive. For example, documents from research-oriented conventions are usually somewhat refereed before acceptance and subsequent submission to ERIC. Such documents certainly receive a greater degree of quality control than a final report which is sent to a funding agency and is almost simultaneously incorporated into ERIC's RIE system.

Table 11 analyzes an arbitrary arrangement of the documents in the present study according to approximations of the degree of quality control before the documents are incorporated into ERIC. The results indicate that on Scales A, C, and E there appear to be differences in quality. A

rank-ordering of the means on these three scales indicates that non-sponsored reports and papers from minor conventions are consistently rated lowest in quality, while the more tightly controlled Educational Testing Service (ETS) papers, the Ph.D. and Masters theses, and AERA papers are rated higher.

The degree of quality control or refereeing would seem to be a factor in these differences. This factor would perhaps be more evident if it could be more adequately isolated. For example, the "Other" category undoubtedly contained some papers which were submitted to large amounts of refereeing which could not be detected in the present study. In addition, conventions, such as that of the AERA, are viewed with varying degrees of formality and informality (cf. Garvey & Griffith, 1971) by different contributors and referees connected with these conventions. Thus quality control probably varies widely within the groups in Table 11. However the present evidence does indicate that considerable differences do in fact exist between the quality control evidently exercised by ETS, for example, and the perhaps more haphazard acceptance policies of non-research-oriented conventions and meetings.

The Decision Setting

Table 12 of Appendix M and Figures 11 through 16 indicate that significant differences in the amount of divergence between Decision Makers and Research Specialists were related to the decision setting. The greatest discrepancy consistently appears in the High Information Grasp--Major Decision setting. Decisions involving major changes are viewed by Stufflebeam, et al. (1971) as necessitating the most intense and most accurate evaluation. It must be noted, however, that the same shortcoming which applied to the attempts to classify the Decision Makers' plans to

implement the research findings apply here. Both information grasp and the magnitude of the change were based on the perceptions of the Decision Maker himself. As was indicated previously, the Decision Makers apparently approached the problem of implementing change from several different points of view, and consequently results based on self-ratings into these categories are somewhat tenuous.

Summary of Results and Discussion

Overall, only 33% of the ERIC documents are backed up by empirical data. (This data is summarized in Table 13). It is certainly possible that some documents could be useful without being supported by empirical data; for example, a good bibliography, a review of the literature, or a discussion of a theory or educational practice could all be useful without empirical data supported by statistical analysis. However, such reports must be viewed primarily as appeals to authority in the absence of supporting data, and thus the fact that 67% of the material disseminated through RIE is without accompanying empirical support is in itself perhaps cause for concern. It is noteworthy that these figures correspond closely to those in the previous Asher and Vockell (1971) pilot study.

The present study, however, has been concerned with the quality of those reports disseminated through RIE which do contain empirical data The evidence presented in this study indicates that, while some high quality reports exist in RIE, the reports are on the whole of low quality. If the quality of the research-oriented documents in RIE is representative of the overall quality of RIE documents, this would mean that only 32.5% of the RIE documents would be rated as "good" or "excellent" quality.

In addition to pointing out the apparent low quality of the research-oriented documents in RIE, this study has provided evidence that this low

quality is in many cases not recognized by the educational practitioner who must make decisions based on the information which he receives. Assuming that these Educational Decision Makers make decisions based on information which they receive (and it is only to the extent that this assumption is true that the information is worth disseminating in the first place), putting low quality information in their hands while they have the impression that it is actually high quality information is likely to lead to implementation in education systems of faulty research.

On the one hand it is apparent that the classification of a report as "excellent" or "incompetent" is a somewhat tenuous endeavor. The purpose of the project on which the report is based, the state of technology in that domain of work, the amount of money available to the researcher or evaluator, the amount of time available for research and/or evaluation, the people badgering the researcher for information--these and many other factors need to be taken into consideration (cf. Runkel & McGrath, 1972, pp. 220 and 427). An evaluation of a project may be useful to a limited audience, even if it is haphazardly performed, simply because a haphazard evaluation may convey at least some new information which they may not have had access to otherwise. Such an evaluation may even be informally communicated to others. But, on the other hand, the transition from a relatively informal to a relatively formal mode of communication has been shown to give an misleading impression of authenticity to such information (Garvey & Griffith, 1971).

The purpose of RIE is to supply rapid, widespread dissemination of educational information. A useful avenue of research would be to determine what savings of time there actually is in this relatively unrefereed process. A relatively slight savings in time would not really be worth a considerable loss in scientific accuracy.

CONCLUSIONS AND RECOMMENDATIONS

Effects of Rapid Dissemination

ERIC's RIE provides a considerable savings in time required to disseminate educational research information. These gains in speed, however, may be accompanied by several undesirable side effects:

(1) The possibility of useful feedback from prepublication exchanges with peers is minimized or eliminated. Garvey and Griffith (1971) have shown that such prepublication exchanges often result in significant modifications in the final version of a manuscript, such as reanalysis of data and redefinition of concepts. Such improvements would make the research much more useful for educational practitioners, but they cannot be expected to occur when it is the prepublication version itself which is being disseminated.

(2) The distinction between formal and informal communication is clouded. Garvey and Griffith (1971) have emphasized the importance of an informal forum where researchers and authors can present half-formed ideas for critical appraisal and commentary from their peers before re-evaluating them and presenting them to the public. Even if such information is somewhat inaccurate it does little harm if it is confined to this informal channel. If such information is disseminated directly to the public, there is evidence that it may lose its identification as informal and be viewed by the public as definitive and sound information.

(3) Inaccurate or low quality information may be disseminated. This is likely to result from the absence of prepublication scrutiny which normally accompanies the refereeing of articles submitted for journal publication. The present study has demonstrated that such low quality of research information is often the case with RIE documents.

(4) The dissemination of large quantities of low quality information in an information channel tends to obfuscate the impact of the high quality information disseminated in the same channel. A comment such as "It's only on ERIC microfiche. It's probably no good anyway," would tend to bias interested decision makers against the good documents which contain high quality research in the same system.

(5) The probability of postpublication scrutiny in the form of critical comments, critiques, rejoinders, and letters to the ERIC Clearinghouse is lowered. In theory, a critique of an RIE document would be accepted and disseminated as rapidly as the original document. Research by Vockell and Asher (1972), however, suggests that active solicitation as a matter of journal policy and practice is required to obtain such critical comments. The fourteen hundred documents reviewed in the present study and in the Asher and Vockell (1970) pilot study included not a single example of a critique of another RIE document. The absence of such postpublication scrutiny eliminates the possibility that a reader who read an erroneous document will later also read a subsequent document where the errors are described and corrected.

(6) The reader has no assurance that the research reported is accurate, of high quality, and related to previous research in the same area. This role of quality control is traditionally allocated to the editors or board of reviewers of a journal. The ill effects resulting

from the absence of refereeing are minimized to the extent that the reader can evaluate the quality of the document himself. However, the present study provides strong evidence that Educational Decision Makers are in many cases unable to make this evaluation. Consequently, there is considerable likelihood that inaccurate, low quality information will be accepted as accurate, high quality information by practitioners who must make decisions based on the research available to them.

The advantages of a rapid dissemination system such as RIE, therefore, must be weighed against these undesirable side effects. A small savings in time is perhaps not worth the risk of encountering these side effects. They can perhaps be reduced by improving and upgrading the system. As it is now, the system might be doing more harm than good for its users.

Rapid Dissemination with Quality Control

An efficient course of action would be to provide maximum savings in time, while simultaneously eliminating the undesirable side effects of rapid dissemination. This would involve changing the RIE system in such a way as to retain as far as possible its rapid dissemination aspects, while also introducing forms of quality control.

The most obvious form of quality control would be prepublication screening: referee the reports and simply decline to disseminate low quality information. High quality documents would be included in RIE immediately. Other documents would be either rejected or returned to the authors for revision. This refereeing could be accomplished either by employing research specialists with each clearinghouse for this purpose, or by sending the documents to outside referees and basing payment on meeting a deadline.

Such refereeing would not, in itself, cause undue delays in publication. The time consuming aspect of most professional publication is not the refereeing process, but rather the space limitation in the journal, which results in a "publication lag." Pay journals are often able to overcome the publication lag by having the authors pay the costs of printing and adding additional pages as necessary. The American Psychological Association journals have a practice of giving the authors the option of paying the extra cost involved in having an article published in an earlier issue of the journal than would be permitted by the publication lag. All other things being equal, refereeing without the publication lag which accompanies page limitations would involve a relatively minor delay.

A source of delay which often accompanies the refereeing process is the revision process. For example, in its initial form a document may contain inadequate descriptions of an experimental program, or the subjects involved in the study, or of the experimental controls. Or perhaps an inappropriate analysis was used. The correction of these inadequacies would involve either a rewriting of the report or performing additional analysis. Such steps would certainly be time consuming and would slow down the speed of dissemination, but it seems obvious that the resulting improvements in quality would make the research more useful to practitioners who would want to use it as a basis for their decision making.

It is also possible, of course, that closer examination and reanalysis would indicate that the document was inaccurate and is not worth reviewing. In this case it could hardly be argued that such a discovery prior to dissemination would be time consuming or otherwise detrimental to the dissemination process. The refereeing process in this case would serve a useful Elimination function.

Such a refereeing process would be most useful to the extent that it is rapid. Likewise, the revision process would be more useful to the extent that it is performed rapidly. Rather than merely returning a document with the assertion that it "might be accepted if revised adequately," many professional journals are adopting a policy of suggesting specific revisions. In other cases the reviewer or editor actually writes a tentative revision and returns this to the authors for approval and further revision prior to publication. Such procedures expedite the review process and could be applied to the RIE review, revision, and dissemination process.

It seems possible that persons who submit reports for dissemination in RIE might differ from persons who submit articles to professional journals in an important respect: journal authors might be to an extent motivated by the emphasis placed on publications as a part of the university promotion process, whereas RIE reports carry less weight in this area. Consequently, authors of RIE reports might be less motivated than authors of journal articles to revise and resubmit a manuscript. This could result in the loss of many desirable manuscripts which would otherwise appear in RIE. Including suggested revisions as a part of the refereeing process could alleviate this problem. If all the author has to do is add some specifically requested information and retype parts of the document, this might not be viewed as a serious burden. When incentives to encourage the original author to revise the document do not exist, in some cases it may even be desirable to have the ERIC clearinghouse revise the document. Precise roles and incentives for the author and the reviewer would have to be refined, but it does seem that the problem is surmountable.

Discussion thus far has focused on prepublication refereeing. Post-publication comments and interactions have been shown to be essential, even in professional journals. A possible application to RIE might be to encourage such interaction by giving priority to very short critiques which point out shortcomings in other RIE documents. Such critiques or rejoinders could be submitted to immediate, high priority review, and quick dissemination could be guaranteed after a factually accurate critique is submitted.

The application of an efficient refereeing process to the RIE dissemination system would deal with the major problems of ERIC emerging from the present study. Such a refereeing process would be similar to the process which accompanies publication in professional journals. The basic premise would not be that only the best examples of professional research would be accepted, but rather that all high quality research would be disseminated as soon as this quality could be established. Moreover, length of the article, which is a major consideration in page-conscious professional journals, would not be a major concern. Lengthy appendices and copies of measurement instruments could be readily included. This is one of the major and unique advantages of RIE. When appropriate, critical comments of these instruments could be added by the author or by the RIE reviewer. Since the refereeing process would be unique to RIE, continuous experimentation, evaluation, and refinement would be necessary. This refereeing and revision process is merely a suggestion. In view of the concern about RIE indicated in the present study, however, such suggestions would seem to provide a useful starting point.

The present report has concerned itself with the research-oriented aspects of RIE. Eighty percent of the documents in RIE contain no

empirical research, and in many cases this absence of empirical data is entirely appropriate. Although it would be rather arbitrary to assume that the quality of this non-research literature is of similarly low quality, it would nevertheless appear to be safe to infer that variations are present in the quality of this literature. To the extent that such variations between high and low quality occur and to the extent that Educational Decision Makers are unlikely to recognize low quality without assistance, the refereeing, revision, and elimination procedures recommended for the research literature in RIE would also apply to the non-research literature.

Areas for Further Research

The findings summarized above suggest several questions for subsequent research:

(1) How much of a publication lag currently exists within the RIE system? RIE documents are accompanied by a "publication date," which indicates when the document was actually written. A comparison of these dates with the actual date the document is indexed and made available through RIE indicates that the median age of a document is about 6-12 months. Many of the documents which appear in RIE may subsequently appear in professional journals. Useful research could be conducted to determine how much longer it takes to publish an article in a professional journal than to release it through RIE.

(2) When the same research is reported in both RIE and in a professional journal, how do the two reports compare in terms of quality? What changes come about as the result of the refereeing the documents undergo when submitted for professional publication?

(3) How does the quality of RIE documents compare to the quality of currently published research journals? Wandt's (1965) study reviewed 1962 journal publications, and it seems likely that the quality of the refereed literature may have changed since that date.

(4) What is the quality of the non-research literature in RIE? In many cases, the absence of empirical data is entirely appropriate; and in such cases the problem is to evaluate this non-research literature by appropriate criteria. In other cases, the information is relatively meaningless or trivial without supporting data. In some cases the authors of the documents apparently believe that they have data when in reality none is present.

(5) Can RIE documents be refereed without a disproportionate loss of time in the dissemination process? Although refereeing of professional journals has been shown to have its own shortcomings, the results of the present study indicate that refereed literature is better than the non-refereed literature. If refereeing can be done quickly and can bring about a degree of quality control, then it would certainly be worth a brief delay which would be involved. The suggestions provided in the previous section would provide a useful starting point for such research.

(6) Can a useful system of postpublication exchange and scrutiny be introduced into RIE? What would be the effects of active solicitation and rapid dissemination of carefully written critiques, critical comments, and rejoinders concerned with information disseminated in RIE (or even elsewhere)?

Conclusion

This study has shown that the research disseminated through RIE is often of low quality and that this low quality is often unrecognized by the Decision Makers who read the research and base decisions upon it. Moreover, it has been shown that Decision Makers who lack research sophistication are most likely to overrate the quality of the reports. It has been argued that such a result is not only a waste of time and money, but rather an active disservice to the Decision Makers.

However, this study has also shown that there are wide variations in the quality of the documents within RIE. Quality in documents seems to be related to the degree of refereeing or quality control prior to its dissemination. Therefore, it would seem useful to introduce a system of rapid refereeing into the RIE system. The implementation of such a system would retain the advantages of RIE's rapid dissemination system, while reducing its negative side effects.

LIST OF REFERENCES

Asher, W. Development, dissemination, and adoption: The need for an elimination function. Journal of Experimental Education, 1972, 40, 14-23.

Asher, W. & Dodson, B. The prediction of automobile accidents following the senior year in high school. Behavioral Research in Highway Safety, 1972, in press.

Asher, W. & V ckell, E. L. Dissemination processes in unrefereed channels (Purdue University, unpublished manuscript, 1970).

Bartos, B. R. A review of instruments developed to be used in the evaluation of the adequacy of reported research. Paper presented at the annual meeting of the American Educational Research Association. February, Los Angeles, 1969.

Baxter, B. (Ed.). Evaluative Research. Pittsburgh: American Institutes for Research, 1970.

Borg, W. R. Educational research: An introduction. New York: David McKay Company, Inc., 1963.

Brady, E. L. & Branscomb, L. M. Information for a changing society. Science, 1972, 175, 961-966.

Campbell, D. T. Reforms as experiments. American Psychologist, 1969, 24, 409-429.

Campbell, D. T. & Stanley, J. C. Experimental and quasi-experimental designs for research. In N. L. Gage (Ed.) Handbook of Research on Teaching. Chicago: Rand McNally and Company, 1963.

Caro, F. G. Issues in the evaluation of social programs. Review of Educational Research, 1971, 41, 87-114.

Clark, K. E. A critical examination of the National Information System for Psychology. American Psychologist, 1971, 26, 325-329.

Corey, S. M. Action research to improve school practices. New York: Teachers College Columbia, 1953.

Dixon, W. J. (Ed.) BMD Biomedical Computer programs. Los Angeles: University of California, 1968.

Educational Resources Information Center. Keeping up. ERIC Clearing House for Education and the Social Sciences, Boulder, Colorado, 1970 (a).

Educational Resources Information Center. Your guide to current educational information. ERIC Clearinghouses, 1970(b).

Farquhar, W. & Krumboltz, J. D. A checklist for evaluating experimental research in psychology and education. Journal of Educational Research, 1959, 52, 354.

Fox, D. J. The research process in education. Chicago: Holt, Rinehart and Winston, 1969.

Garvey, W. D. & Griffith, B. C. The American Psychological Association's Project on Scientific Information Exchange in Psychology. Journal of Counseling Psychology, 1963, 10, 297-302.

Garvey, W. D. & Griffith, B. C. Scientific information exchange in psychology. Science, 1964, 146, 1655-1659(a).

Garvey, W. D. & Griffith, B. C. The structure, objectives, and findings of a study of scientific information exchange in psychology. American Documentation, 1964, 15, 258-267(b).

Garvey, W. D. & Griffith, B. C. Scientific communication: The dissemination system in psychology and a theoretical framework for planning innovations. American Psychologist, 1965, 20, 157-164.

Garvey, W. D. & Griffith, B. C. Studies in social innovations in scientific communication in psychology. American Psychologist, 1966, 21, 1019-1036.

Garvey, W. D. & Griffith, B. C. Scientific communication as a social system. Science, 1967, 157, 1011-1016.

Garvey, W. D. & Griffith, B. C. Scientific communication: Its role in the conduct of research and creation of knowledge. American Psychologist, 1971, 26, 349-362.

Garvey, W. D., Lin, N., & Nelson, C. E. Communication in the physical and the social sciences. Science, 1970, 170, 1166-1173.

Gephart, W. J. & Bartos, B. B. Profiling instructional package. Bloomington: Phi Delta Kappa Research Service Center, 1969.

Gideonse, H. D. An output-oriented model of research and development and its relationship to educational improvement. Journal of Experimental Education, 1968, 37, 157-163.

Glass, G. V. Educational piltdown men. Phi Delta Kappan, 1968, 50, 148-151.

Grobman, H. Evaluation of Curriculum Projects. AERA Monograph Series on Curriculum Evaluation. Chicago: Rand McNally and Company, 1968.

Guba, E. G., & Clark, D. L. An examination of potential change roles in education. Paper read at Symposium on Innovation in Planning School Curricula held by National Education Association Committee for Study of Instruction, Airielhouse, Virginia, October, 1965.

Guilford, J. P. Fundamental statistics in psychology and education. New York: McGraw-Hill, 1964.

Havelock, R. G. Planning for innovation through dissemination and utilization of knowledge. Center for Research on Utilization of Scientific Knowledge, Institute for Social Research, University of Michigan, Ann Arbor, Michigan, 1969.

Holt, R. R. Some neglected assumptions and problems in psychology's information crisis. American Psychologist, 1971, 26, 331-334.

Kerlinger, F. N. Foundations of behavioral research. Chicago: Holt, Rinehart & Winston, 1964.

Kohr, R. L. & Suydam, M. N. An instrument for evaluating survey research. Journal of Educational Research, 1970, 61, 78-85.

Kronick, D. A. Information Media. Encyclopedia Britannica. Vol. 12, 244D-244I, 1969.

Licklider, J. C. R. A crux in scientific and technical communication. American Psychologist, 1966, 21, 1044-1051.

Loevinger, J. Monographs (letter to the editor). APA Monitor, 1972, 3 (1 & 2), 9.

Mann, J. Evaluating educational programs: A symposium. The Urban Review, 1969, 3(4), 12-23.

McReynolds, P. Reliability ratings of research papers. American Psychologist, 1971, 26, 400-401.

Michael, W. B. Teacher personnel: A brief evaluation of the research reviewed. Review of Educational Research, 1963, 33, 443.

Mouly, G. J. The science of educational research. New York: American Book Company, 1963.

National Academy of Sciences Committee on research in the Life Sciences of The Committee on Science and Public Policy. Communication in the Life Sciences, The Life Sciences, Washington, D. C.: National Academy of Sciences, 1970, 405-427.

Nunnally, J. C. Psychometric theory. New York: McGraw-Hill, 1967.

Persell, C. H. Research Evaluation instrument. Columbia University, Unpublished, 1966.

Riedesel, C. A. A change in "focus". The Arithmetic Teacher, 1971, 18, 113-114.

Rummel, J. F. An introduction to research procedures in education. (2d. ed.) New York: Harper and Row, Publishers, 1964.

Runkel, P. & McGrath, J. E. Research on human behavior: A systematic guide to method. Chicago: Holt, Rinehart & Winston, 1972.

Scriven, M. The philosophy of science in educational research. Review of Educational Research, 1960, 30, 422-429.

Scriven, M. The methodology of evaluation. In B. Othanel Smith (Ed.) Perspectives of Curriculum Evaluation. Chicago: Rand McNally, 1967.

Sieber, S. D. Survey research in education: The Case of the misconstrued technique. Phi Delta Kappan, 1968 (Jan.), 49, 273-276.

Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Guba, E. G., Hammond, R. L., Marriman, H. O., Provus, M. M. Educational Evaluation and Decision Making. Itasca, Illinois: Teacock, 1971.

Suchman, E. A. Evaluative research: Principles and practice in public service and social action programs. New York: Russell Sage Foundation, 1967.

Suydam, Marilyn N. An instrument for evaluating experimental educational research reports. Journal of Educational Research, 1968, 61, 200-203.

Travers, R. M. W. An introduction to educational research. (3d ed.) New York: Macmillan, 1969.

Van Dalen, D. B. Understanding educational research. New York: McGraw-Hill, 1966.

Veldman, D. J. Fortran programming forthe behavioral sciences. Chicago: Holt, Rinehart & Winston, 1967.

Vockell, E. L., & Asher, W. Critiques and critical comments in educational research journals. Educational Researcher, 1972, 1(5), 11-12.

Wandt, E. Cross-section of educational research. New York: David McKay Company, Inc. 1965.

Wandt, E., et al. An evaluation of educational research in published journals. Report of the Committee on evaluation of research, American Educational Research Association, 1967.

Ziman, J. M. Information, communication, knowledge. Nature, 1969, 224, 318-324.

APPENDIX A

THE DOCUMENTS IN THE STUDY

ED 042 087 Influences of the Kentucky cooperative extension home economics program. Part I

ED 042 149 Influence of initial and final consonants on vowel duration in CVC syllables.

ED 042 213 Proprioceptive factors in operative timing.

ED 042 287 A 3-year study of continuity of creative growth under a cognitive-structured approach to educational stimulation.

ED 042 363 Optimal use of a computer based instruction system in an existing urban school district. Final Report.

ED 042 413 Report on the subcommittee of the status of academic women on the Berkeley campus.

ED 042 499 The modification of teacher behaviors which modify child behaviors. Progress Report.

ED 042 513 The "Tell-and-Find Picture Games" for young children.

ED 042 563 Language survey of entering primary students (rural schools)

ED 042 587 Transfer from word components to words and vice versa - beginning reading.

ED 042 763 Pupil motivation: A rewarding experience.

ED 042 863 A look at combination class effects at Emerson Elementary School.

ED 043 049 Evaluation of a program for developing creative thinking in teachers and children at the fourth and fifth grade levels.

ED 043 099 National assessment of educational progress. Summary of Report 1. Science: National Results, July 1970.

ED 043 287 Some characteristics of finalists in the 1966 national achievement scholarship program.

ED 043 313 Attitude change of freshman college students toward their role as raters of teacher behavior.

ED 043 449 Staffing, equipping, and operating public school reading clinics.

ED 043 463 Blueprint for a successful paraprofessional tutorial program.

ED 043 587 Preparing teachers to involve parents in children's learning. Project Report.

ED 043 599 Evaluation report of cooperative urban teacher education program. Fall 1969, Working Paper, Vo. 3 No. 1.

ED 044 063 The MLA Foreign Language Proficiency Tests for Teachers: Advanced Students: Analysis of the performance of native speakers: Comparison with that of NDEA Summer Institute Participants.

ED 044 087 A survey of rank, salary, promotion, and tenure policies in 15 colleges and universities.

ED 044 099 Student and instructor expectations of community college instructors.

ED 044 387 Educational innovation: Research and evaluation techniques. (Rep. of 1967 Summer Institutes)

ED 044 499 Competencies needed by seed production and distribution company employees.

ED 044 649 Evaluation of adult basic education workshop in western Tennessee, Dec. 14, 1960.

ED 044 749 Counselor bias and the female occupational role.

ED 044 863 The effects of a program of learning games upon selected academic abilities in children with learning difficulties.

ED 044 963 EPDA Institute in German for cooperating teachers and teacher trainers. Final Report.

ED 045 049 Student personality characteristics and optimal college learning conditions. Final Report.

ED 045 463 Evaluation report: Project Insight 1969-1970.

ED 045 763 A research project of the North Carolina Advancement School. Summer Session 1968.

ED 046 049 Bases of classification of geometric concepts used by children of varying characteristics. Report from the Project on Situation Variables and Efficiency of Concept Learning. Part I.

ED 046 149 A follow-up study of graduates from a vocational rehabilitation program in a residential training center for the mentally retarded.

ED 046 249 Aptitude-treatment interactions in two studies of learning from lecture instruction.

ED 046 499 Piaget's concept of classification: A comparative study of socially disadvantaged and middle-class young children.

ED 046 513 Developmental groupings of pre-school children.

ED 046 563 Oregon state migrant education program. Annual Project Report and Evaluation 1967-68.

ED 046 599 An inter-disciplinary evaluation of a summer program for the rural disadvantaged youth in Nelson County, Virginia. Final Report.

ED 046 987 Levels of educational performance and related factors in Michigan.

ED 046 999 Clinical-diagnostic use of the WPPSI in predicting learning disabilities in grade one.

ED 047 099 Teaching beliefs and classroom methodology of selected project and cooperative plan distributive education teacher coordinators.

ED 047 113 Follow-up of 1965 graduates of Wisconsin Schools of Vocational, Technical, and Adult Education.

ED 047 413 Preparing communication consultants for a school district: A study of some effects of laboratory training.

ED 047 487 Mathematics curriculum innovation and evaluation for educable mentally retarded. Final Report.

ED 047 513 The interaction of associative memory and general reasoning with availability and complexity of examples in a computer-assisted instruction task.

ED 047 699 Reaching the unreached; report of the regional action workshop devoted to the "non-user" of library services.

ED 047 899 Predictive values of selected auditory perceptual factors in relation to measured first-grade reading achievement.

ED 047 987 The Oakland County mathematics project.

ED 048 099 Pupil control idealogy and teacher influence in the classroom.

ED 048 349 Assessment of teachers' attitudes toward an individualized approach to reading instruction.

ED 048 487 Testing of the guidance program, project able; development and evaluation of an experimental curriculum for the New Quincy Massachusetts Vocational Technical School.

ED 048 549 Models of diffusing knowledge to community college faculty.

ED 048 613 Client motivation and rehabilitation counseling outcome.

ED 048 799 Auditory discrimination and Identification in foreign language learning.

ED 048 863 Cost analysis: First step toward differentiated funding.

ED 048 899 Study and comparison of the indexing of the "bibliography of agriculture" in relation to the indexing of 15 other secondary services.

ED 049 287 An investigation of selected procedures for the development and evaluation of hierarchical curriculum structures.

ED 049 313 Using the MMPI to predict teacher behavior.

ED 049 387 Development of educational programs for new careers in recreation services for the disabled.

ED 049 449 Inventory of continuing education activities in Pennsylvania institutes of higher learning for the academic year 1969-70. Final Report.

ED 049 713 Encouraging convergence of opinion, through the use of the Delphi Technique, in the process of identifying an institutional goal.

ED 049 813 Correlation of paired-associate performance with school achievement as a function of task and sample variation.

ED 050 013 The effects of perceived teacher attitudes on student achievement.

ED 050 087 Effects of commonly known meanings on determining obscure meanings of multiple-meaning words in context. Report for the project on individually guided elementary language arts.

ED 050 149 The mini rat: Its development and some evidence of its validity.

ED 050 263 SUM (service unit management): An organizational approach to improved patient care.

ED 050 399 Manifest characteristics of interactive sequencing in the classroom.

ED 050 463 Organizational climate: Another dimension to the process of innovation.

ED 050 663 A 4-year survey of DePaul University's School of Education master degree graduates for 1966 thru 1969. First Interim Report.

ED 050 687 Evaluation and selection of academic interns: 1967-68.

ED 050 863 The nongraded-trimester-minicourse concept: Report to date.

ED 051 049 A study of the measured readability level of selected intermediate grade social studies textbooks.

ED 051 087 An evaluation of the program "supportive training for inexperienced and new teachers" (STINT) in New York City schools.

ED 051 099 The Stanford secondary teacher education program 1959-69: A preliminary analysis of graduate career patterns.

ED 051 113 Human interaction: A behavioral objectives curriculum in human relations.

ED 051 249 Relationship between teacher' vocabulary usage and the vocabulary of kindergarten and first grade students. Final Report.

ED 051 749 The transfer student: An institutional variable in the analysis of junior year college marks.

ED 051 863 The effect of indexing specificity on retrieval performance.

ED 052 213 Attitudinal correlates of children's speech characteristics. Final Report.

ED 052 249 Effect of the response format on the differential measurement of traits in the Thorndike Dimensions of Temperamental Personality Inventory.

ED 052 299 A feasibility study in determining individual practice profiles of physicians as a basis for continuing education of these physicians utilizing a postgraduate preceptor technique.

ED 052 499 The development of cognitive performance criteria for use by secondary school teachers. Final Report.

ED 052 713 Black and nonblack youth: Finances and college attendance.

ED 052 749 An investigation of entrance characteristics related to types of college dropouts. Final Report.

ED 052 863 Academic performance and school integration: A multi-ethnic analysis.

ED 053 063 The study of student behavior in science as a result of modification of certain identifiable teacher behaviors. Final Report.

ED 053 099 Assessment of the impact of the teacher corps program at the University of Miami and participating schools in Southern Florida.

ED 053 113 A no-grammar approach to sentence power: John C. Mellon's sentence-combining games.

ED 053 163 Interaction between college effects and students' aptitudes.

ED 053 187 The genetic components of verbal divergent thinking and short term memory.

ED 053 349 Management practices of cotton producers in Lauderdale County, Tennessee.

ED 053 413 Washington, D.C. drug seminar for school personnel and students; Aug. 31-Sept. 4, 1970. Evaluation Report.

ED 053 549 The development of a measure of state epistemic curiosity.

ED 053 649 Estimating parents' contribution to college costs. The accuracy of three measures of succeeding year family net income.

ED 053 713 Class size on community college economic education.

ED 053 849 The organizational climate in schools having high concentrations of Mexican-Americans.

ED 053 887 Model Programs: Reading Yuba County Reading-Learning Center, Marysville, California.

ED 054 187 Research in art education: The development of perception in art production of kindergarten students. Final Report.

ED 054 387 Survey for comparison of effectiveness of home management with residency versus home management without residency laboratory 1966-68. Final Report.

ED 053 987 Development of a more flexible physical science laboratory program for non-science majors with superior high school science backgrounds. Final Report.

ED 054 287 Communicative dimensions of mother-child interaction as they affect the self esteem of the child.

APPENDIX B

LETTER FOR INITIAL CONTACT WITH EDUCATIONAL DECISION MAKERS

This letter was used in only about 40% of the cases. In the other cases, initial contact was made by phone. The phone message was similar to the content of this letter.

PURDUE UNIVERSITY
SCHOOL OF HUMANITIES, SOCIAL SCIENCE AND EDUCATION
LAFAYETTE, INDIANA 47907

DEPARTMENT OF EDUCATION

Dear Colleague:

We are undertaking a study of the quality and usefulness of the research information disseminated through Research in Education, an abstracting service of ERIC (Educational Resources Information Center).

We need educational decision makers to read and evaluate these research reports. We are able to pay a token honorarium of $15.00 for a review of a short paper (which would take about one hour to review) and up to $30.00 for a longer report (which would take perhaps two hours to review). We shall attempt to match the reports with your areas of interest.

We shall ask you to review one of the reports and return your analysis to us within a week to 10 days or so. If you are able to participate, we would appreciate it if you would return the enclosed forms now in the self addressed envelope.

Thank you for your cooperation.

Sincerely,

William Asher
Professor of Education and
Psychological Sciences

WA:pw
Enc.

APPENDIX C

THE DECISION MAKERS

Stuart A. Anderson
Professor of Administration
Sangamon State University

Jane Bebb
Director of Pupil Services
School City of Mishawaka

Frederick B. Bechtold
Assistant Superintendent of Instruction
Marion Community Schools
Marion, Indiana

Larry L. Beckley
Principal
Crete Elementary School
Crete, Illinois

Lawrence Beymer
Professor of Education and Assistant Dean
Indiana State University

Mark Breitenstein
Language Arts Coordinator
Our Lady of Lourdes Elementary School
Cincinnati, Ohio

Donn Brolin
Director, Undergraduate Training in Vocational Rehabilitation
University of Wisconsin

James E. Brumbaugh
Education Consultant
Indiana State Department of Public Instruction

James E. Brumbaugh
Education Consultant
Indiana State Department of Public Instruction

Douglas Bryant
Professor of Education
North Carolina State University

Napoleon Bryant, Jr.
Assistant Professor of Education
Xavier University

Naomi S. Clark
Teacher
Bond Hill Elementary School

James Clemmons
Director of Reading
Forest Hills School District

Louise Cuyler
Project Director, Head Start and Child Development Programs of the Greater Cincinnati Area
Community Action Commission

Harriet Darrow
Director of Continuing Education for Women
Indiana State University

Richard Dickson
Adult Program Coordinator
Wabash Center
Lafayette, Indiana

Frank DiSilvestro
Educational Consultant
Indiana State Department of Public Instruction

Robert A. Eaverly
Principal
Sayre Junior High School
Philadelphia, Pa.

Kenneth W. Ehle
Consultant in Mathematics
Indiana State Department of Public Instruction

John L. Eick
Assistant Professor of French
Xavier University

Lee Elliott
Assistant Professor of Psychology
Xavier University

Charles F. Elton
Professor of Higher Education
University of Kentucky

Edward Eustice
Director of Graduate Studies
Olivet Nazarene College

Joseph L. Felix
Associate Director for Research
and Evaluation
Cincinnati Public Schools

Sister M. Florianne
Director, School of Nursing
St. Elizabeth Hospital School
of Nursing
Lafayette, Indiana

Lloyd E. Frohreich
Assistant Professor of Educational
Administration
University of Wisconsin

James Gaffney
Assistant Dean of Graduate School
Xavier University

Robert G. George
Coordinator of Instructional systems
Technology
Indiana State University

John Gerard
General Manager, Seed Grain Division
Syler, Inc.
Plymouth, Indiana

Dave Gilman
Professor of Education
Indiana State University

Melvin Goldberg
Research Associate
Institute for Educationa Development

A. C. Gondring
Retired Elementary Principal
and Teacher
Perry Central Community School
Corporation
Leopold, Indiana

William J. Hamrick
Assistant Superintendent for
Instruction
Vigo County School Corporation
Terre Haute, Indiana

Beverly Hankenhoff
Practical Arts and Continuing
Education Director
New Castle Area Vocational
School and Chrysler High School
New Castle, Indiana

Robert H. Helmes
Dean, College of Continuing
Education and Summer Sessions
Xavier University

Marjorie W. Hesler
Assistant Professor of Speech
Communication
Indiana State University

Michael D. Hinkle
Principal
Joseph F. Tuttle Jr. High
Crawfordsville, Indiana

Philip J. Hobbs
Assistant Chairman, Division of
Education
Indiana-Purdue Regional Center
Indianapolis, Indiana

James Hogan
Director of Curriculum
Forest Hills School District
Cincinnati, Ohio

Donald Huff
Assistant Director, Student Affairs
Indiana Vocational Technical
Institute
Lafayette, Indiana

Charles A. Hunter
Coordinator, Sociological Research
Dallas Independent School District

John A. Hutchinson
Curriculum Consultant
Indiana State Department of
Public Instruction

Jim Jacobs
Director, Research and Development
Cincinnati Public Schools

Rudolf Jacobsen
Principal
Cruft School
Terre Haute, Indiana

Frank W. Jerse
Professor of Educational Psychology
Indiana State University

Simon O. Jiboku
Social Studies Supervisor
Indiana State Department of
Public Instruction

Bruce C. Landis
Coordinator for Teacher Education
Indiana State Department of
Public Instruction

Conrad Lane
Supervisor of Student Teachers
Ball State University

Harold W. Lang
Director, State-Wide Testing
Dallas Independent School District

Jene M. Galvin
Co-Founder
New Morning Community School
Cincinnati, Ohio

Harley Lautenschlager
Director of Laboratory School
Indiana State University

Terry O. Lowe
Consultant for Counseling and
Guidance Services
Indiana State Department of
Public Instruction

David H. McGaw
Biology Program Director
West Lafayette High School
West Lafayette, Indiana

C. Eugene McLemore
Director of Admissions
Hanover College
Hanover, Indiana

Frank W. Medley
Adjunct Associate Professor of
Foreign Languages
Indiana State University

Mescal Messmore
Language Arts Consultant
Indiana State Department of
Public Instruction

Leon Met
Project Coordinator, Student
Leadership Development Center
Cincinnati Public Schools

Joel Milgram
Assistant Professor of Education
University of Cincinnati

Rudolph P. Minneman
Director of Elementary Education
Tippecanoe School Corporation
Lafayette, Indiana

Jim E. Montgomery
School Library Consultant
Indiana State Department of
Public Instruction

Don Moore
Art Consultant
Indiana State Department of
Public Instruction

Marjorie Morehouse
Remedial Reading Teacher
Lafayette School Corporation
Lafayette, Indiana

Stanley R. Morgan
Assistant Superintendant
Salt Lake City Schools
Salt Lake City, Utah

Stephen D. Mueller
Guidance Counselor
North College Hill Junior-Senior
High School
Cincinnati, Ohio

George C. Nagle
Vice Principal
Lake Oswego Elementary School
Lake Oswego, Oregon

Earl E. Nelson
Director of Elementary Education
Lafayette School Corporation
Lafayette, Indiana

Henry W. Nelson
Dean of Students
Wheaton College
Wheaton, Illinois

Patricia O'Connor
Research Associate
School of Dentistry
University of Michigan

George H. Olson
Consultant on Evaluation
Dallas Independent School District

James B. O'Neill
Assistant Director
Center of Economics Education
University of Delaware

William R. Osmon
Dean, Student Administrative Services
Indiana State University

Frederick K. Packer
Director, Instructional Materials Center
Coos County Intermediate Education District
North Bend, Oregon

Barbara Pashos
Reading Consultant
Indiana State Department of Public Instruction

John P. Pohlman
Director, Lodge Learning Laboratory
Xavier University

Harbison Pool
Project Director, Project ORBIT
Oberlin City School District
Oberlin, Ohio

Alton V. Potts
Regional Director
Tippewa Technical Institute
Lafayette, Indiana

Gerald L. Quatman
Professor of Psychology
Xavier University

Henry F. Raichle
Director of Research
Pinellas County School System
Clearwater, Florida

James Roland
Director of Special Education
Urbana Public Schools
Urbana, Illinois

Herbert L. Ross
Associate Professor
School of Business
Indiana State University

John Rothwell
Assistant to Dean - Assistant Professor of Economics
Xavier University

JoAnne S. Rowe
Instructor in Home Economics
Indiana State University

Kenneth F. Scheurer
Chairman Department of Education
Xavier University

Roman J. Schweikert
Associate Professor of Education
Xavier University

Dorothy June Sciarra
Assistant Professor of Child Development
University of Cincinnati

Bernard T. Shirk
English Supervisor
Gary Schools
Gary, Indiana

Annie Sims
Director
Lower Price Hill Community Day Center
Cincinnati, Ohio

Dennis H. Sorge
Federal Programs Evaluator
Scottsdale Public Schools
Phoenix, Arizona

Walter L. Sullins
Director, Center for Educational Research
Indiana State University

Carl T. Swift
Superintendent
Pike County School Corporation
Petersburg, Indiana

Tom Tegarden
Coordinator of Supportive Services for Adult Education
Indiana State Department of Public Instruction

Alice Trese
Montessori Pre-School Instructor
St. Augustine's School
Cincinnati, Ohio

Christine Ann Trusler
Elementary Teacher
Northwest Local School District
Cincinnati, Ohio

Mark Van Horn
Drug and Narcotic Consultant
Indiana State Department of Public Instruction

John William VanKerk
Language Laboratory Director
Indiana State University

Gerald L. Verland
Division of Program Research and Design
Cincinnati Public Schools

Ivan D. Wagner
Director of Planning Evaluation
Indiana State Department of Public Instruction

E. Dewey Wahl
Staff Director
Industrial Management
Purdue University

William J. Webster
Director, Research and Evaluation
Dallas Independent School District

Mary Wine
Welfare Rights Coordinator
Santa Maria Neighborhood House
Cincinnati, Ohio

Thelma M. Wootton
Director
Wells Memorial Library
Lafayette, Indiana

Robert E. Wubbolding
Adjunct Associate Professor
Guidance Department
Xavier University

Stanley M. Yabroff
District Associate
Institute for Educational Development

Marshall A. Moore
Director of Secondary Guidance
Richmond Community Schools
Richmond, Indiana

APPENDIX D

THE INITIAL LETTER TO THE RESEARCH SPECIALISTS

This letter was sent to as many as possible of the research specialists in Wandt's (1967) study.

PURDUE UNIVERSITY
SCHOOL OF HUMANITIES, SOCIAL SCIENCE AND EDUCATION
LAFAYETTE, INDIANA 47907

DEPARTMENT OF EDUCATION

March 1, 1972

Dear

In 1962, we participated in an AERA study to examine the quality of published educational research literature. That study was under the direction of David Ryans and Edwin Wandt, and was reported by AERA as, "An Evaluation of Educational Research in published Journals" (mimeo, 1967), and in Wandt's Cross-Section of Educational Research (1965).

Edward Vockell and I have recently received funding for a U.S.O.E. Small Grant to undertake a similar study of the unpublished research disseminated through Research in Education, the abstracting service of ERIC (Educational Resources Information Center).

We would like to have the same set of reviewers as the 1962 AERA study. However, this time we can pay a token honorarium of $15.00 for a review of a short report (which would take about one hour to review) and up to $30.00 for a longer report (which would take perhaps 2 hours to review).

We will ask you to review one of the reports and return your analysis of its quality within a week to 10 days or so. We would appreciate it if you would return the enclosed post-card now indicating your availability. We shall again attempt to match the reports with your areas of interest.

Thank you for your cooperation.

Sincerely,

William Asher
Professor of Education and
Psychological Sciences

WA:pw
Enc.

APPENDIX E

THE EDUCATIONAL RESEARCH SPECIALISTS

Gale R. Adkins*
Harold H. Abelson
Kenneth E. Anderson
William Asher
Nancy Bayley
Carl Bereiter
John E. Bicknell
Walter R. Borg
Russel W. Burris
William R. Carricker
John B. Carroll
Werrett Wallace Charters
David L. Clark
Thomas D. Clemons
John F. Cogswell
Desmond L. Cook*
Francis G. Cornell
John T. Dailey
Henry S. Dyer
Ruth E. Eckert
Max D. Englehart
Daniel D. Feder
Donald W. Felker
John C. Flanagan
Benjamin Fruchter
N. L. Gage
George L. Gropper
Egon G. Guba
Elizabeth Hagen
Ray M. Hall
Berj Harootunian
Dale B. Harris
Albert Hieronymous
John L. Holland
Kai Jensen
Felix F. Kapstein
Patricia M. Kay
Evan R. Keisler
Bert Y. Kersh
David J. Klaus
Dorothy M. Knoell
Russell P. Kropp
Philip Lambert
Charles R. Langmuir
Gerald V. Lannholm
D. Welty Lefever
Roger T. Lennon
Elizabeth Steiner Maccia
James B. Macdonald
Walter H. MacGintie
Samuel T. Mayo*
Ernest McDaniel
Wilbert J. McKeachie
John D. McLeil
Harold E. Metzel
Carol Miller
William C. Morse
George J. Mouly
Claude L. Nemzek
Victor H. Noll
David B. Orr
R. Travis Osborne
Ellis B. Page
Herman J. Peters *
Beeman N. Phillips*
Douglas Porter
Thomas J. Quirk
William Rabinowitz
William Reitz
James H. Ricks
Helen M. Robinson
Herman Roemmich
Gale W. Rose
Philip J. Runkel*
Einar R. Ryden
Gilbert Sax*
Dale P. Scannell
Alice Yeomans Scates
Richard E. Schutz
C. Winfield Scott*
Saul B. Sells
Robert S. Soars
Seth Spaulding
William H. Stegeman
George G. Stern
Lawrence Stolurow
Kalmer P. Stordahl*
Donald Thistlewaite
George G. Thompson
Robert M. W. Travers
Arthur E. Traxler
William W. Turnbull
Richard Turner
R. Keith Van Wagenen
Donald J. Veldman
Norman E. Wallen
Alexander G. Wesman
Merlin C. Wittrock
Asahel D. Woodruff
J. Wayne Wrightstone

* Read a second document

APPENDIX F

COVER LETTER SENT TO DECISION MAKERS WITH DOCUMENT TO BE REVIEWED

The instructions were in this letter and on the Information Quality and Uses Form. Further instructions were provided only upon request.

PURDUE UNIVERSITY
SCHOOL OF HUMANITIES, SOCIAL SCIENCE AND EDUCATION
LAFAYETTE INDIANA 47907

DEPARTMENT OF EDUCATION

Dear Colleague:

Enclosed is the research report which you recently agreed to review as part of our study on educational research dissemination. Please follow the following steps:

1. Complete the "General Information Form" (pink sheet).

2. Fill in the "Familiarity with ERIC Questionnaire" (yellow sheet). Do not return to this form after you have read the enclosed research report.

3. Read the enclosed report to evaluate it. Use the "Information Quality and Uses" questionnaire (white paper).

4. Complete the "Project Certification" form (white paper).

5. Return all four of the above in the enclosed stamped envelope. (You can keep the research report.) The university will mail you a check for your honorarium sometime after we receive the above information.

We have attempted to match the reports with your areas of interest. In some cases we may have failed in this attempt. If you feel totally unable to review the report you have received, return it to us and we'll try to make some adjustment.

If you have any questions, please contact us.

Thank you for your participation in this study.

Sincerely,

William Asher
Professor of Education and
Psychological Sciences

WA:pw
Enc.

APPENDIX G

SAMPLE OF FOLLOW-UP TELEGRAM

Recently you received a report to read and evaluate as part of our "Information Quality and Educational Decision Making" study. We would like to receive all the completed evaluation forms by July 28. If this is impossible, please either return the document to us or phone us collect (317-749-2845). We will reimburse you for any additional expenses you incur. If your records indicate that you have already returned your evaluation form, or if you have any other questions, please call us collect immediately.

William Asher

APPENDIX H

COVER LETTER SENT TO RESEARCH SPECIALISTS WITH DOCUMENT TO BE REVIEWED

The instructions were in this letter and on the Form for Evaluating Quality of Research. Further instructions were provided only upon request.

PURDUE UNIVERSITY
SCHOOL OF HUMANITIES, SOCIAL SCIENCE AND EDUCATION
LAFAYETTE INDIANA 47907

DEPARTMENT OF EDUCATION

Enclosed is the research report which you recently agreed to review as part of our study on "Information Quality and Educational Decision Making."

Please read the report and evaluate it by filling in the "Form for Evaluating Quality of Research." (white paper)

We have also enclosed copies of several more detailed instruments for assessing the quality of educational research. These are included for your own use. They may be helpful to you in calling your attention to important points which should be kept in mind while you are filling out the other evaluation forms. Do not return any of these forms on colored paper.

Return the completed evaluation form in the enclosed stamped envelope. Our business office also requests that you fill in the "Project Certification" form so that we can mail you your honorarium promptly.

We have attempted to match the reports with your areas of interest In some cases we may have failed in this attempt. If you feel totally unable to review the report you have received, return it to us and we'll try to make some adjustment.

If you have any questions, please contact us.

Thank you for your participation in our study.

Sincerely,

William Asher
Professor of Education and
Psychological Sciences

WA:pw
Enc.

APPENDIX I

Information Quality and Uses Form.

Instructions: Please fill in the information as indicated. You are encouraged to make use of any additional references which may be available to you. If you need additional space, feel free to write on the other side of the pages or to add additional pages.

ERIC #_____ Reviewer_____

Brief Title _____

Author _____

1. Brief summary: (Problem; subjects--number and description; treatments or classifications; procedures; analyses; results)

2. Educational significance (Does the report address itself to an important educational problem? Does it have a sound theoretical basis? What is the relation to known research?)

3. Research Design (Are there any problems in the design of the research or methodology employed?)

4. Rate the quality of this research according to one of the following criteria:

_____1. Completely incompetent--A "horrible example."

_____2. Poor -- Some serious defects. A

_____3. Mediocre -- Not good, not bad.

_____4. Good -- A few minor defects.

_____5. Excellent -- A model of good practice.

5. If you were the editor of a journal to which this report was submitted for publication, would you: B

_____Accept -- Accept the article and publish it without major revision.

_____Revise -- Ask for revisions and publish it if the revisions were made.

_____Reject -- Reject the article.

6. Estimate the approximate point where you would place yourself on the continuum below with regard to knowledge of the subject covered by the report.

1	2	3	4	5	6	7
know very little						know a great deal

7. Give your estimate of how significant or important you think a change involving the subject of this report would be, if implemented in your school or school system.

1	2	3	4	5	6	7
very minor change						very major change

8. What recommendations would you make with regard to the implementation of the information contained in this report?

RATING FORM

Article Number: _____ Rater's Name: _____

On the basis of an over-all impression of the research article you have just read, please rate it (using the scale shown below on the left side of the page) as to both substance and methodology, checking one box in each column.

		(A) SUBSTANCE		(B) METHODOLOGY
		The ideas or empirical findings presented; their substantive contribution to any field(s) of theoretical knowledge in education or a discipline	The ideas or empirical findings presented; their substantive contribution to any field(s) of educational practice	The study's utilization of (or contribution to) research methods
		Rating Scale C	Rating Scale D	Rating Scale E
(I)	Ranks with the best empirical research studies known to me; on a par with the top 5 or 10 percent in this respect.	(1)	(1)	(1)
(II)	Better than average, though not "outstanding" in this respect.	(2)	(2)	(2)
(III)	Run of the mill in this respect; neither better nor poorer than the bulk of research that I have seen.	(3)	(3)	(3)
(IV)	Not up to "average" standards; "less-than-mediocre" in this respect (although not altogether lacking).	(4)	(4)	(4)
(V)	Incompetent in this respect: among the poorest examples of "research" that I have encountered.	(5)	(5)	(5)

APPENDIX J

Form for Evaluating Quality of Research.

Instructions: Please fill in the information as indicated. You are encouraged to make use of any additional checklists and references which may be available to you. If you need additional space, feel free to write on the other side of the pages or to add additional pages.

Are you currently teaching a course in research design or research methodology, or have you ever taught such a course?

_____Yes _____No

Please list here the approximate number of research reports you have either published in professional journals or presented at conventions or professional meetings:_____

ERIC #_____ Reviewer_____

Brief Title _____

Author _____

1. Brief summary: (Problem; subjects--number and description; treatments or classifications; procedures; analyses; results)

2. Educational significance (Does the report address itself to an important educational problem? Does it have a sound theoretical basis? What is the relation to known research?)

3. Research Design (Are the problems, objectives, procedures, and the relationships among all three clearly and logically stated? Does the statement of procedures include, where applicable, information on sampling techniques, controls, data gathered, instruments used, and statistical and other analyses made?)

4. Rate the quality of this research according to one of the following criteria:

A

____5. Excellent--A model of good practice

____4. Good -- A few minor defects

____3. Mediocre -- Not good, not bad

____2. Poor -- Some serious defects

____1. Completely Incompetent -- A "horrible example"

5. If you were the editor of a journal to which this report was submitted for publication, would you:

B

____Accept -- Accept the article and publish it without revision

____Revise -- Ask for revisions and publish if the revisions were made

____Reject -- Reject the article

RATING FORM

Article Number: _____ Rater's Name: _____

On the basis of an over-all impression of the research article you have just read, please rate it (using the scale shown below on the left side of the page) as to both substance and methodology, checking one box in each column.

		(A) SUBSTANCE		(B) METHODOLOGY
		The ideas or empirical findings presented; their substantive contribution to any field(s) of theoretical knowledge in education or a discipline	The ideas or empirical findings presented; their substantive contribution to any field(s) of educational practice	The study's utilization of (or contribution to) research methods
		Rating Scale C	Rating Scale D	Rating Scale E
(I)	Ranks with the best empirical research studies known to me; on a par with the top 5 or 10 percent in this respect.	(1)	(1)	(1)
(II)	Better than average, though not "outstanding" in this respect.	(2)	(2)	(2)
(III)	Run of the mill in this respect; neither better nor poorer than the bulk of research that I have seen.	(3)	(3)	(3)
(IV)	Not up to "average" standards; "less-than-mediocre" in this respect (although not altogether lacking).	(4)	(4)	(4)
(V)	Incompetent in this respect: among the poorest examples of "research" that I have encountered.	(5)	(5)	(5)

APPENDIX K

SUPPLEMENTARY EVALUATION CHECKLISTS

SENT TO RESEARCH SPECIALISTS

These were not returned and are not a part of the analysis.

EVALUATING EDUCATIONAL RESEARCH ARTICLES (Edwin Wandt)

Rate each characteristic below on a 5-point scale as follows:

5 = Excellent; a model of good practice
4 = Good; a few minor defects
3 = Mediocre; not good, not bad
2 = Poor; some serious defects
1 = Completely incompetent; a "horrible example"

1 _____ Problem is clearly stated.

2 _____ Hypotheses are clearly stated.

3 _____ Problem is significant.

4 _____ Assumptions are clearly stated.

5 _____ Limitations of the study are stated.

6 _____ Important terms are defined.

7 _____ Relationship of the problem to previous research is made clear.

8 _____ Research design is described fully.

9 _____ Research design is appropriate to the solution of the problem.

10 _____ Research design is free of specific weaknesses.

11 _____ Method of sampling is appropriate.

12 _____ Population and sample are fully described.

13 _____ Data-gathering methods or procedures are described.

14 _____ Data-gathering methods or procedures are appropriate to the solution of the problem.

15 _____ Data-gathering methods or procedures are utilized correctly.

16 _____ Validity and reliability of the evidence gathered are established.

17 _____ Appropriate methods are selected to analyze the data.

18 _____ Methods utilized in analyzing the data are applied correctly.

19 _____ Results of the analysis are presented clearly.

20 _____ Conclusions are clearly stated.

21 _____ Conclusions are substantiated by the evidence presented.

22 _____ Generalizations are confined to the population from which the sample was drawn.

23 _____ Report is clearly written.

24 _____ Report is logically organized.

25 _____ Tone of the report displays an unbiased, impartial scientific attitude.

26 _____ Conclusions are significant.

27 _____ Conclusions are relevant to the problem.

AN INSTRUMENT FOR EVALUATING EXPERIMENTAL EDUCATIONAL RESEARCH REPORTS

Marilyn N. Suydam
The Pennsylvania State University

Directions:

Evaluate with the nine italicized questions which follow. The quality of the article in terms of each question should be rated on a five-point scale. The specifications for these five points are:

1) Excellent: all requirements for the question are met; nothing essential could be added
2) Very good: most requirements are met
3) Good: some requirements are met
4) Fair: a few requirements are met
5) Poor: none or too few of the requirements are met

Certain "key points" should be considered in ascertaining a rating for each question. These are listed below the question, followed by adjectives which indicate the continuum on which the "key point" should be assessed. Do NOT make a response to these "key points". They are intended to focus the attention of all raters on the same pertinent aspects of each question.

Please make only nine responses for each article, one for each question.

1. How practically or theoretically significant is the problem? (1-2-3-4-5)
   a. Purpose (important-non-important)
   b. Problem origin
      1) Rationale (logical-illogical)
      2) Previous research (appropriate-inappropriate)

2. How clearly defined is the problem? (1-2-3-4-5)
   a. Question (operational-vague)
   b. Hypothesis (es) (relevant-irrelevant) (logical-illogical)
   c. Independent variable(s) (operational-vague) (relevant-irrelevant)
   d. Dependent variable(s) (operational-vague) (relevant-irrelevant)

3. How well does the design answer the research questions? (1-2-3-4-5)
   a. Paradigm (appropriate-inappropriate)
   b. Hypothesis(es) (testable-untestable)
   c. Procedures (clear-unclear)
   d. Treatments (replicable-unreplicable) (appropriate-inappropriate)
   e. Duration (appropriate-inappropriate)

4. How adequately does the design control variables? (1-2-3-4-5)
   a. Independent variable(s) (uncontaminated-contaminated)
   b. Administration of treatment (rigorous-unrigorous)
   c. Teacher or group factors (controlled-uncontrolled)
   d. Subject or experimenter bias (controlled-uncontrolled)
   e. Halo effect (controlled-uncontrolled)
   f. Extraneous factors (controlled-uncontrolled)
   g. Individual factors (controlled-uncontrolled)

5. How properly is the sample selected for the design and purpose of the research? (1-2-3-4-5)
   a. Population (appropriate-inappropriate)
   b. Drawing of sample (random-unspecified)
   c. Assignment of treatment (random-unspecified)
   d. Size (appropriate-inappropriate)
   e. Characteristics (appropriate-inappropriate)

6. How valid and reliable are the measuring instruments or observational techniques? (1-2-3-4-5)
   a. Instrument or technique
      1) Description (excellent-poor)
      2) Validity (appropriate-inappropriate)
      3) Reliability for population (excellent-poor)
   b. Procedures of data collection (careful-careless)

7. How valid are the techniques of analysis of data? (1-2-3-4-5)
   a. Statistical tests
      1) Basic assumptions (satisfied-unclear)
      2) Relation to design (appropriate-inappropriate)
   b. Data
      1) Treatment (appropriate-inappropriate)
      2) Presentation (clear-unclear)
      3) Level of significance (specified-unspecified)
         (appropriate-inappropriate)
      4) Discussion (accurate-inaccurate)

8. How appropriate are the interpretations and generalizations from the data? (1-2-3-4-5)
   a. Consistency with results (excellent-poor)
   b. Generalizations (reasonable-exaggerated)
   c. Implications (reasonable-exaggerated)
   d. Limitations (noted-not noted)

9. How adequately is the research reported? (1-2-3-4-5)
   a. Organization (excellent-poor)
   b. Style (clear-vague)
   c. Grammar (good-poor)
   d. Completeness (excellent-poor)
   (replicable-unreplicable)

INSTRUMENT FOR EVALUATING SURVEY RESEARCH REPORTS

(Richard L. Kohr, The Pennsylvania State University)

Directions: The following instrument is to be used for evaluating survey research reports within the framework of curriculum research. It is composed of nine major questions which are underlined. You are to rate the quality of the report in terms of each of these nine questions using the following 5-point scale:

(1) Excellent: all requirements for the question are met; nothing essential could be added.
(2) Very good: most requirements are met.
(3) Good: some requirements are met.
(4) Fair: a few requirements are met.
(5) Poor: none or too few of the requirements are met,

In determining a rating for each question, certain "key points" should be considered. These are listed below the question, followed by adjectives which indicate the continuum on which the key point should be assessed. Do NOT make a response to these "key points." They are intended to focus the attention of all raters on the same pertinent aspects of each question. In some studies certain "key points" may be irrelevant. In such cases base your judgment on such "key points" as are relevant. It is also possible that you may think of "key points" not included among those listed under a major question. Where relevant, such additional "key points" may be used in assessing that question. There may be some instances in which none of the "key points" seem relevant or where the report fails to supply sufficient information. If this occurs, evaluate the report in terms of what you think should have been done and/or what information should have been included.

Please make only nine responses for each article, one for each question.

1 _____ How practically or theoretically significant is the problem? (1-2-3-4-5)
- a. Purpose (important - non-important)
- b. Problem origin
  - (1) Rationale (logical - illogical
  - (2) Previous research (related - unrelated)
- c. Generalizability (extensive - limited)

2 _____ How clearly defined is the survey problem? (1-2-3-4-5)
- a. Objectives and procedures (specified - unspecified) (operational - vague)
- b. Delimitations (noted - not noted)
- c. Variables
  - (1) Control (relevant - irrelevant) (operational - vague)
  - (2) Dependent (relevant - irrelevant) (operational - vague)

3 _____ How relevant and how well defined is the population?
- a. Precise definition of population
  - (1) Geographical limits (specified - unspecified)
  - (2) Time period covered (specified - unspecified)
  - (3) Sociological description (specified - unsepcified)
  - (4) Sampling units (specified - unsepecified)
- b. Relevance of defined population to problem (relevant - irrelevant)

4 ____ How adequate are the sampling procedures? (1-2-3-4-5)
- a. Adequacy of sampling frame
  - (1) Time period covered (current - outdated)
  - (2) Inclusiveness of defined population (complete - incomplete)
- b. Method of sampling (specified - unspecified)
- c. Obtained sample (appropriate - inappropriate)
  - (1) Size (sufficient - insufficient)
  - (2) Representativeness (adequate - inadequate)

5 ____ How adequately are sources of error controlled? (1-2-3-4-5)
- a. Sampling error (controlled - uncontrolled)
- b. Non-response (controlled - uncontrolled)
- c. Interviewer bias (controlled - uncontrolled)
- d. Response error (controlled - uncontrolled)
- e. Response set (controlled - uncontrolled)
- f. Experimenter bias (controlled - uncontrolled)
- g. Teacher effect (controlled - uncontrolled)
- h. Control variables (controlled - uncontrolled)
- i. Extraneous factors (controlled - uncontrolled)
- j. Qualifications of research personnel (Interviewers, etc.) (controlled - uncontrolled)

6 ____ How adequate are the measuring instruments? (1-2-3-4-5)
- a. Choice of measurement technique(s) (appropriate - inappropriate)
- b. Instrument(s)
  - (1) Development of instrument (pretested - not pretested)
  - (2) Description of administration and scoring procedures (clear - unclear)
  - (3) Wording of statements or questions (clear - unclear)
  - (4) Sequence of statements or questions (logical - illogical) (random - fixed)
  - (5) Evidence of reliability (appropriate - inappropriate)
  - (6) Evidence of validity (appropriate - inappropriate) (satisfactory - unsatisfactory)
- c. Rules for categorizing (specified - unspecified)

7 ____ How appropriate is the statistical analysis of the data?
- a. Procedures of data collection (specified - unspecified) (careful - careless)
- b. Relation of obtained data to objectives (essential - unessential) (sufficient - insufficient)
- c. Descriptive measures
  - (1) Statistics (appropriate - inappropriate)
  - (2) Evaluation of descriptive data (appropriate - inappropriate)
  - (3) Establishment of relationships (appropriate - inappr opriate)
- d. Statistical tests
  - (1) Basic assumptions (satisfied - unsatisfied)
  - (2) Relation to procedures (appropriate - inappropriate)
  - (3) Significance levels (specified - unspecified)
- e. Description of results (accurate - inaccurate)

8 ____ How reasonable are the conclusions drawn from the data?

a. Interpretations (consistent - inconsistent)
(reasonable - exaggerated)

b. Generalizations (reasonable - exaggerated)

c. Implications (reasonable - exaggerated)

d. Qualifications

(1) Discussion of methodological problems and errors (comprehensive - limited)

(2) Alternative explanations (noted - not noted)

(3) Other limitations (noted - not noted)

9 ____ How adequately is the research reported? (1-2-3-4-5)

a. Organization (excellent - poor)

b. Style (clear - vague)

c. Grammar and mechanics (excellent - poor)

d. Completeness (replicable - unreplicable)
(complete - incomplete)

e. Presentation of statistics (clear - unclear)

APPENDIX L

General Information Form

If you are available to participate in this research project, please fill in the following information and answer the brief items in the attached questionnaire.

Name_____
Address_____
_____

Phone_____

Employment or school affiliation:

School:_____
_____

Position: (specific)_____
_____

Type of school:

_____K-6

_____Junior high school

_____High school

_____College

_____(other) Specify:_____
_____

Number of years as a teacher and/or administrator:_____yrs.

Degrees held:

_____less than bachelor's

_____Bachelor's

_____Master's

_____Doctorate

_____(other) Specify:_____

Other relevant information: (Area of interest, etc.)

APPENDIX M

TABLES AND FIGURES

This appendix contains the tables and figures referred to in the Results section of this report. Several notes are useful in interpreting these tables and figures:

(1) The directionality of rating is reversed on Rating Scale A. This is true both for the ratings of the Research Specialists and those of the Decision Makers. A high score on Scale A indicates high quality. On all other scales, a high score indicates low quality.

(2) The term "Group" (G) refers to a group of documents.

(3) The term "Trial" (T) refers to the repeated measurements on the same documents by both a Research Specialist (trial 1) and by a Decision Maker (trial 2).

TABLE 2

RELIABILITIES OF THE RATING SCALES*

Research Specialists

Rating Scale A (Overall Rating)	.56	(9)
Rating Scale B (Accept-Revise-Reject)	.15	(10)
Rating Scale C (Theoretical Contribution)	.30	(11)
Rating Scale D (Practical Contribution)	.51	(11)
Rating Scale E (Methodology)	.27	(11)

Decision Makers

Rating Scale A (Overall Rating)	.36	(8)
Rating Scale B (Accept-Revise-Reject)	.24	(8)
Rating Scale C (Theoretical Contribution)	.51	(7)
Rating Scale D (Practical Contribution)	.37	(7)
Rating Scale E (Methodology)	.48	(7)

*Figures in parentheses after reliability coefficients indicate the number of documents included in the computations.

TABLE 3

INTERCORRELATIONS OF THE RATING SCALES

	Research Specialists					Decision Makers				
	*A	B	C	D	E	*A	B	C	D	E
*A	1.00									
B	-.63	1.00								
C	-.73	.44	1.00							
D	-.73	.53	.63	1.00						
E	-.76	.54	.72	.63	1.00					
*A	.04	-.00	-.06	-.11	-.08	1.00				
B	-.06	.12	.10	.10	.13	-.69	1.00			
C	-.09	.10	.14	.14	.18	-.77	.63	1.00		
D	-.00	.02	.04	.14	.09	-.74	.71	.80	1.00	
E	.04	-.09	.06	.05	.08	-.77	.58	.75	.70	1.00

*The directionality of the rating is reversed on Scale A. High scores on Scale A indicate a rating of high quality, whereas on all other scales high scores indicate low quality.

TABLE 4

COMPARATIVE RATINGS OF RESEARCH SPECIALISTS

AND DECISION MAKERS ON RATING SCALES

Rating Scale A. (Overall Rating)

	Research Specialists		Decision Makers	
	No.	Pct.	No.	Pct.
5. Excellent	5	5.2	15	16.3
4. Good	26	27.3	32	34.7
3. Mediocre	22	23.1	29	31.5
2. Poor	32	33.6	15	16.3
1. Completely incompetent	10	10.5	1	10
* No response	7		10	

$\chi^2 = 19.99$ (df = 4) $p < .001$

Rating Scale B. (Accept-Revise-Reject)

	No.	Pct.	No.	Pct.
1. Accept	8	8.2	30	32.6
2. Revise	37	38.1	39	42.3
3. Reject	52	53.6	23	25.0
* No response	5		10	

$\chi^2 = 23.87$ (df = 2) $p < .001$

Rating Scale C. (Theoretical Contribution)

	No.	Pct.	No.	Pct.
1. Excellent	2	2.1	7	7.6
2. Above average	13	13.6	28	30.7
3. Mediocre	27	28.4	37	40.6
4. Below average	37	38.9	11	12.0
5. Incompetent	16	16.8	8	8.7
* No response	7		10	

$\chi^2 = 26.47$ (DF = 4) $p < .001$

TABLE 4, contd

Rating Scale D. (Practical Contribution)

	Research Specialists		Decision Makers	
	No.	Pct.	No.	Pct.
1. Excellent	3	3.0	11	11.9
2. Above average	21	21.4	39	42.3
3. Mediocre	34	34.6	23	25.0
4. Below average	28	28.5	15	16.3
5. Incompetent	12	12.2	4	4.3
* No response	4		10	

$\chi^2 = 22.20$ (df = 4) $p < .001$

Rating Scale E. (Methodology)

	No.	Pct.	No.	Pct.
1. Excellent	7	7.2	9	9.7
2. Above average	16	16.4	14	15.2
3. Mediocre	24	24.7	41	44.5
4. Below average	28	28.8	19	20.6
5. Incompetent	22	22.6	9	9.7
* No response	5		10	

$\chi^2 = 11.81$ (df = 4) $p < .05$

*"No response" includes evaluations not returned and blank responses on the rating scale.

TABLE 5

MEANS AND STANDARD DEVIATIONS ON THE RATING SCALES

	Mean	Standard Deviation
Research Specialists		
Rating Scale A	*2.83	1.10
Rating Scale B	1.43	.66
Rating Scale C	3.54	.99
Rating Scale D	3.25	1.02
Rating Scale E	3.43	1.21
Decision Makers		
Rating Scale A	*3.48	.98
Rating Scale B	1.92	.75
Rating Scale C	2.83	1.03
Rating Scale D	2.58	1.03
Rating Scale E	3.05	1.07

*The directionality of the rating is reversed on scale A. High scores on this scale indicate a rating of high quality, whereas on all other scales a higher score indicates lower quality.

TABLE 6

COMPARISON OF THE RESEARCH SPECIALISTS RATINGS ON THE ACCEPT-REVISE-REJECT SCALE WITH THOSE IN WANDT'S (1967) STUDY

	n	$\bar{X}$	S. D.
Wandt (1967)	125	2.21	.74
Present Study	*96	2.43	.66

$t = 2.29$

$p < .05$

* Several Research Specialists left this item blank.

TABLE 7

RATINGS OF RESEARCH SPECIALISTS COMPARED TO RATINGS OF DECISION MAKERS

ANALYSIS FOR VARIABLE 1 (Rating Scale A)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.2325	171		
BETWEEN	1.1678	85		
TRIALS*	21.6337	1	20.462	.0001
ERROR (T)	1.0573	85		
N VALID	86			

T MEAN	1	2
	2.8023	3.5116

ANALYSIS FOR VARIABLE 2 (Rating Scale B)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	.5765	175		
BETWEEN	.5734	87		
TRIALS*	12.0227	1	26.836	.0000
ERROR (T)	.4480	87		
N VALID	88			

T MEAN	1	2
	2.4318	1.9091

TABLE 7, cont.

ANALYSIS FOR VARIABLE 3 (Rating Scale C)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.1654	169		
BETWEEN	1.1482	84		
TRIALS*	28.0059	1	32.451	.0000
ERROR (T)	.8630	84		
N VALID	85			

T MEAN	1	2
	3.6235	2.8118

ANALYSIS FOR VARIABLE 4 (Rating Scale D)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.1932	177		
BETWEEN	1.2067	88		
TRIALS*	25.9775	1	28.929	.0000
ERROR (T)	.8980	88		
N VALID	89			

T MEAN	1	2
	3.3146	2.5506

TABLE 7, cont.

ANALYSIS FOR VARIABLE 5 (Rating Scale E)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.3397	175		
BETWEEN	1.4017	87		
TRIALS*	9.5511	1	8.071	.0057
ERROR (T)	1.1833	87		
N VALID	88			

T MEAN	1	2
	3.5000	3.0341

* Note that "Trial" (T) refers to the repeated measurements on the same document by the Research Specialists (trial 1) and by the Decision Makers (trial 2).

TABLE 8

ANALYSIS OF VARIANCE BY EDUCATIONAL LEVEL OF DECISION MAKERS (PH. D. VS. NON-PH. D.)

GROUP 1 44 SUBJECTS. Documents Read by Ph. D. Decision Makers

GROUP 2 54 SUBJECTS. Documents Read by Non-Ph. D Decision Makers

ANALYSIS FOR VARIABLE 1 (Rating Scale A)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.2350	169		
BETWEEN	1.1811	84		
GROUPS	.1616	1	.135	.7149
ERROR (G)	1.1934	83		
WITHIN	1.2082	85		
TRIALS *	20.4765	1	19.621	.0001
G BY T	2.4026	1	2.302	.1291
ERROR (T)	1.0436	83		

G MEAN	1	2
	3.1923	3.1304
N VALID	39	46

T MEAN	1	2
	2.8118	3.5059

G BY T	1	2
1	2.9744	3.4103
2	2.6739	3.5870

TABLE 8, cont.

ANALYSIS FOR VARIABLE ? (Rating Scale B)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	.5713	173		
BETWEEN	.5794	86		
GROUPS	1.1112	1	1.939	.1639
ERROR (G)	.5731	85		
WITHIN	.5632	87		
TRIALS*	11.1264	1	25.132	.0000
G BY T	.2420	1	.547	.5315
ERROR (T)	.4427	85		

G MEAN	1	2
	2.2632	2.1020
N VALID	38	49

T MEAN	1	2
	2.4253	1.9195

G BY T	1	2
1	2.4737	2.0526
2	2.3878	1.8163

TABLE 8, cont.

ANALYSIS FOR VARIABLE 3 (Rating Scale C)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.1668	167		
BETWEEN	1.1609	83		
GROUPS	1.4640	1	1.265	.2630
ERROR (G)	1.1572	82		
WITHIN	1.1726	84		
TRIALS *	26.7202	1	30.909	.0000
G BY T	.8926	1	1.033	.3135
ERROR (T)	.8645	82		

G MEAN	1	2
	3.3205	3.1333
N VALID	39	45

T MEAN	1	2
	3.6190	2.8214

G BY T	1	2
1	3.6410	3.0000
2	3.6000	2.6667

TABLE 8, cont.

ANALYSIS FOR VARIABLE 4 (Rating Scale D)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.1774	175		
BETWEEN	1.2131	87		
GROUPS	.8002	1	.657	.5747
ERROR (G)	1.2179	86		
WITHIN	1.1420	88		
TRIALS *	24.0057	1	28.111	.0000
G BY T	3.0547	1	3.577	.0587
ERROR (T)	.8539	86		

G MEAN	1	2
	3.0000	2.8646
N VALID	40	48

T MEAN	1	2
	3.2955	2.5568

G BY T	1	2
1	3.2250	2.7750
2	3.3542	2.3750

TABLE 8, cont.

ANALYSIS FOR VARIABLE 5 (Rating Scale E)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.3454	173		
BETWEEN	1.4042	86		
GROUPS	1.1562	1	.822	.6296
ERROR (G)	1.4071	85		
WITHIN	1.2874	87		
TRIALS*	9.1954	1	7.757	.0067
G BY T	2.0405	1	1.721	.1900
ERROR (T)	1.1855	85		

G MEAN	1	2
	3.1875	3.3511
N VALID	40	47

T MEAN	1	2
	3.5057	3.0460

G BY T	1	2
1	3.3000	3.0750
2	3.6809	3.0213

* Note that "Trial" (T) refers to the repeated measurements on the same document by the Research Specialists (trial 1) and by the Decision Makers (trial 2).

TABLE 9

ANALYSIS OF VARIANCE BY INSTITUTIONAL AFFILIATION OF DECISION MAKER

ANALYSIS FOR VARIABLE 1 (Rating Scale A)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.2192	169		
BETWEEN	1.1494	84		
GROUPS	1.6973	3	1.503	.2187
ERROR (G)	1.1291	81		
WITHIN	1.2882	85		
TRIALS *	20.4765	1	20.368	.0001
G BY T	2.5302	3	2.517	.0628
ERROR (T)	1.0053	81		

G MEAN	1	2	3	4
	3.0000	3.1765	3.3500	2.8571
N VALID	7	34	30	14

T MEAN	1	2
	2.8235	3.5176

G BY T	1	2	
1	3.0000	3.0000	Research Organizations
2	2.9118	3.4412	Colleges & Universities
3	3.0333	3.6667	State & Local Systems
4	2.0714	3.6429	Individual Schools

TABLE 9, cont.

ANALYSIS FOR VARIABLE ? (Rating Scale B)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	.5790	173		
BETWEEN	.5775	86		
GROUPS	.1250	3	.211	.8890
ERROR (G)	.5939	83		
WITHIN	.5805	87		
TRIALS *	11.6379	1	26.744	.0000
G BY T	.9147	3	2.102	.1048
ERROR (T)	.4352	83		

G MEAN	1	2	3	4
	2.2143	2.2188	2.1212	2.1333
N VALID	7	32	33	15

T MEAN	1	2
	2.4253	1.9080

G BY T	1	2	
1	2.1429	2.2857	Research Organizations
2	2.4063	2.0313	Colleges & Universities
3	2.4848	1.7576	State & Local Systems
4	2.4667	1.8000	Individual Schools

TABLE 9, cont.

ANALYSIS FOR VARIABLE 3 (Rating Scale C)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.1599	167		
BETWEEN	1.1471	83		
GROUPS	.8821	3	.762	.5213
ERROR (G)	1.1570	80		
WITHIN	1.1726	84		
TRIALS *	26.7202	1	33.047	.0000
G BY T	2.3649	3	2.925	.0380
ERROR (T)	.8086	80		

G MEAN	1	2	3	4
	3.5714	3.2500	3.1333	3.1000
N VALID	7	32	30	15

T MEAN	1	2
	3.6071	2.8095

G BY T	1	2	
1	3.4286	3.7143	Research Organizations
2	3.5938	2.9063	Colleges & Universities
3	3.5667	2.7000	State & Local Systems
4	3.8000	2.4000	Individual Schools

TABLE 9, cont.

ANALYSIS FOR VARIABLE 4 (Rating Scale D)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.1774	175		
BETWEEN	1.2131	87		
GROUPS	.0459	3	.037	.9900
ERROR (G)	1.2548	84		
WITHIN	1.1420	88		
TRIALS*	24.0057	1	30.312	.0000
G BY T	3.3232	3	4.196	.0083
ERROR (T)	.7920	84		

G MEAN	1	2	3	4
	3.0000	2.9412	2.9063	2.9000
N VALID	7	34	32	15

T MEAN	1	2	
	3.2955	2.5568	
G BY T	1	2	
1	2.7143	3.2857	Research Organizations
2	3.2353	2.6471	Colleges and Universities
3	3.3438	2.4688	State and Local Systems
4	3.6000	2.2000	Individual Schools

TABLE 9, cont.

ANALYSIS FOR VARIABLE 5 (Rating Scale E)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.3285	173		
BETWEEN	1.3702	86		
GROUPS	.5262	3	.376	.7739
ERROR (G)	1.4007	83		
WITHIN	1.2874	87		
TRIALS*	10.1379	1	8.801	.0042
G BY T	2.0842	3	1.809	.1504
ERROR (T)	1.1519	83		

G MEAN	1	2	3	4
	3.0714	3.2794	3.2097	3.4333
N VALID	7	34	31	15

T MEAN	1	2
	3.5057	3.0230

G BY T	1	2	
1	2.8571	3.2857	Research Organizations
2	3.4706	3.0882	Colleges & Universities
3	3.4516	2.9677	State & Local Systems
4	4.0000	2.8667	Individual Schools

* Note that "Trials" (T) refers to the repeated measurements on the same document by the Research Specialists (trial 1) and by the Decision Makers (trial 2).

TABLE 10

ANALYSIS OF VARIANCE BY CLEARINGHOUSE

GROUP 1	5 SUBJECTS.
GROUP 2	8 SUBJECTS.
GROUP 3	4 SUBJECTS.
GROUP 4	4 SUBJECTS.
GROUP 5	12 SUBJECTS.
GROUP 6	5 SUBJECTS.
GROUP 7	6 SUBJECTS.
GROUP 8	5 SUBJECTS.
GROUP 9	10 SUBJECTS.
GROUP10	6 SUBJECTS.
GROUP11	9 SUBJECTS.
GROUP12	8 SUBJECTS.

ANALYSIS FOR VARIABLE 1

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.2560	76		
GROUPS	1.2156	11	.963	.5108
ERROR (G)	1.2628	65		

G MEAN

1	2	3	4
2.4000	2.8750	2.2500	3.2500
5	6	7	8
3.0000	2.4000	2.6667	1.7500
9	10	11	12
2.8000	2.7500	3.5000	3.1250

TABLE 10, cont.

ANALYSIS FOR VARIABLE 2

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	.4309	77		
GROUPS	.6508	11	1.651	.1049
ERROR (G)	.3942	66		

G MEAN

1	2	3	4
2.8000	2.3750	2.5000	2.0000
5	6	7	8
2.6364	2.8000	2.4000	2.5000
9	10	11	12
2.8000	2.4000	2.0000	2.0000

ANALYSIS FOR VARIABLE 3

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	.9566	76		
GROUPS	1.1716	11	1.273	.2594
ERROR (G)	.9202	65		

G MEAN

1	2	3	4
4.2000	3.3750	4.0000	2.5000
5	6	7	8
3.7273	3.8000	4.0000	4.2500
9	10	11	12
3.5556	3.4000	3.2222	3.2857

TABLE 10, cont.

ANALYSIS FOR VARIABLE 4

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.0500	78		
GROUPS	1.0266	11	.974	.5212
ERROR (G)	1.0538	67		

G MEAN

1	2	3	4
3.8000	3.3750	4.0000	3.0000
5	6	7	8
3.0909	3.6000	3.1667	4.0000
9	10	11	12
3.1000	2.6000	2.8889	3.0000

ANALYSIS FOR VARIABLE 5

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.4725	77		
GROUPS	2.0937	11	1.529	.1420
ERROR (G)	1.3690	66		

G MEAN

1	2	3	4
4.2000	3.0000	4.2500	2.0000
5	6	7	8
3.1818	4.2000	3.3333	4.2500
9	10	11	12
3.7000	3.6000	3.2222	3.2857

TABLE 11

ANALYSIS OF VARIANCE BY SPONSORSHIP OF REPORTS

GROUP 1	11 SUBJECTS.	AERA Convention Papers
GROUP 2	9 SUBJECTS.	Smaller conventions and minor journals
GROUP 3	4 SUBJECTS.	Educational Testing Service
GROUP 4	5 SUBJECTS.	Masters or Doctoral Theses
GROUP 5	73 SUBJECTS.	Other

ANALYSIS FOR VARIABLE 1 (Rating Scale A)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.2267	94		
GROUPS	2.8097	4	2.430	.0526
ERROR (G)	1.1563	90		

G MEAN	1	2	3	4	5
	3.3000	1.8571	3.2500	3.4000	2.7971
N VALID	10	7	4	5	69

ANALYSIS FOR VARIABLE 2 (Rating Scale B)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	.4356	96		
GROUPS	.3303	4	.750	.5626
ERROR (G)	.4401	92		

G MEAN	1	2	3	4	5
	2.2727	2.6667	2.7500	2.2500	2.4203
N VALID	11	9	4	4	69

TABLE 11, cont.

ANALYSIS FOR VARIABLE 3 (Rating Scale C)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	.9951	94		
GROUPS	3.1561	4	3.511	.0105
ERROR (G)	.8990	90		

G MEAN	1	2	3	4	5
	3.3636	4.5000	2.5000	3.2000	3.5522
N VALID	11	8	4	5	67

ANALYSIS FOR VARIABLE 4 (Rating Scale D)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.0580	97		
GROUPS	.7865	4	.735	.5727
ERROR (G)	1.0696	93		

G MEAN	1	2	3	4	5
	3.0000	3.5556	2.7500	3.6000	3.2609
N VALID	11	9	4	5	69

ANALYSIS FOR VARIABLE 5 (Rating Scale E)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.4588	96		
GROUPS	3.1512	4	2.275	.0662
ERROR (G)	1.3852	92		

G MEAN	1	2	3	4	5
	3.3636	4.0000	2.2500	2.6000	3.5294
N VALID	11	9	4	5	68

TABLE 12

ANALYSIS OF VARIANCE BY DECISION SITUATION

ANALYSIS FOR VARIABLE 1 (Rating Scale A)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.1861	141		
BETWEEN	1.0177	70		
GROUPS	.3803	3	.364	.7824
ERROR (G)	1.0462	67		
WITHIN	1.3521	71		
TRIALS*	19.0423	1	19.789	.0001
G BY T	4.1619	3	4.325	.0078
ERROR (T)	.9623	67		

G MEAN	1	2	3	4
	3.3043	3.1765	3.0909	3.1000
N VALID	23	17	11	20

T MEAN	1	2
	2.8169	3.5493

G BY T	1	2	Information	Change
1 (n=25)	2.5217	4.0870	High	Major
2 (n=19)	3.1176	3.2353	High	Minor
3 (n=11)	2.8182	3.3636	Low	Major
4 (n=24)	2.9000	3.3000	Low	Minor

TABLE 12, cont.

ANALYSIS FOR VARIABLE 2 (Rating Scale B)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	.5988	147		
BETWEEN	.0099	73		
GROUPS	.5922	3	.970	.5864
ERROR (G)	.6106	70		
WITHIN	.5878	74		
TRIALS*	10.2770	1	25.781	.0000
G BY T	1.7729	3	4.447	.0067
ERROR (T)	.3986	70		

G MEAN	1	2	3	4
	2.0400	2.1176	2.0909	2.3095
N VALID	25	17	11	21

T MEAN	1	2
	2.4054	1.8784

G BY T	1	2	Information	Change
1 (n=25)	2.5200	1.5600	High	Major
2 (n=19)	2.1176	2.1176	High	Minor
3 (n=11)	2.4545	1.7273	Low	Major
4 (n=24)	2.4762	2.1429	Low	Minor

TABLE 12, cont.

ANALYSIS FOR VARIABLE 3 (Rating Scale C)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.1040	141		
BETWEEN	1.0380	70		
GROUPS	1.2453	3	1.210	.3124
ERROR (G)	1.0287	67		
WITHIN	1.1690	71		
TRIALS *	23.6901	1	33.565	.0000
G BY T	4.0073	3	5.678	.0019
ERROR (T)	.7058	67		

G MEAN	1	2	3	4
	3.0000	3.4000	3.2500	3.3095
N VALID	25	15	10	21

T MEAN	1	2
	3.6197	2.8028

G BY T	1	2	Information	Change
1 (n=25)	3.8000	2.2000	High	Major
2 (n=19)	3.5333	3.2667	High	Minor
3 (n=11)	3.5000	3.0000	Low	Major
4 (n=24)	3.5238	3.0952	Low	Minor

TABLE 12, cont.

ANALYSIS FOR VARIABLE 4 (Rating Scale D)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.1724	149		
BETWEEN	1.1445	74		
GROUPS	1.7822	3	1.595	.1970
ERROR (G)	1.1176	71		
WITHIN	1.2000	75		
TRIALS *	22.4267	1	30.054	.0000
G BY T	4.8638	3	6.518	.0009
ERROR (T)	.7462	71		

G MEAN	1	2	3	4
	2.6600	2.9167	3.1364	3.0714
N VALID	25	18	11	21

T MEAN	1	2
	3.2933	2.5200

G BY T	1	2	Information	Change
1 (n=25)	3.4000	1.9200	High	Major
2 (n=19)	2.8889	2.9444	High	Minor
3 (n=11)	3.7273	2.5455	Low	Major
4 (n=24)	3.2857	2.8571	Low	Minor

TABLE 12, cont.

ANALYSIS FOR VARIABLE 5 (Rating Scale E)

SOURCE	MEAN SQUARE	D.F.	F-RATIO	P
TOTAL	1.2670	147		
BETWEEN	1.2088	73		
GROUPS	1.2054	3	.997	.5993
ERROR (G)	1.2090	70		
WITHIN	1.3243	74		
TRIALS *	8.7568	1	8.234	.0056
G BY T	4.9339	3	4.640	.0054
ERROR (T)	1.0634	70		

G MEAN	1	2	3	4
	3.0400	3.3824	3.3182	3.3810
N VALID	25	17	11	21

T MEAN	1	2
	3.5000	3.0135

G BY T	1	2	Information	Change
1 (n=25)	3.7200	2.3600	High	Major
2 (n=19)	3.3529	3.4118	High	Minor
3 (n=11)	3.4545	3.1818	Low	Major
4 (n=24)	3.3810	3.3810	Low	Minor

*Note that "Trial" (T) refers to the repeated measurements on the same document by the Research Specialists (trial 1) and by the Decision Makers (trial 2).

TABLE 13

SUMMARY OF THE CHARACTERISTICS OF THE ERIC DOCUMENTS

	1970* (Vockell & Asher, 1970)	1971 (Present Study)
Total Documents Examined	745	617
Number of Documents with Research	155	203
Percent of Documents with Research	20.8	32.9
Percent of Research Documents Rated Good or Excellent	39.3	32.5
Percent of Total Documents Supported by Good or Excellent Research	6.7	5.0

* The 1970 sample was based on the abstracts from the first ten months of that year. Since this was a pilot study, the criteria for classification were not as carefully supervised as in the present study. Each document was classified only once by one of sixteen reviewers. No reliability check was attempted during the pilot study.

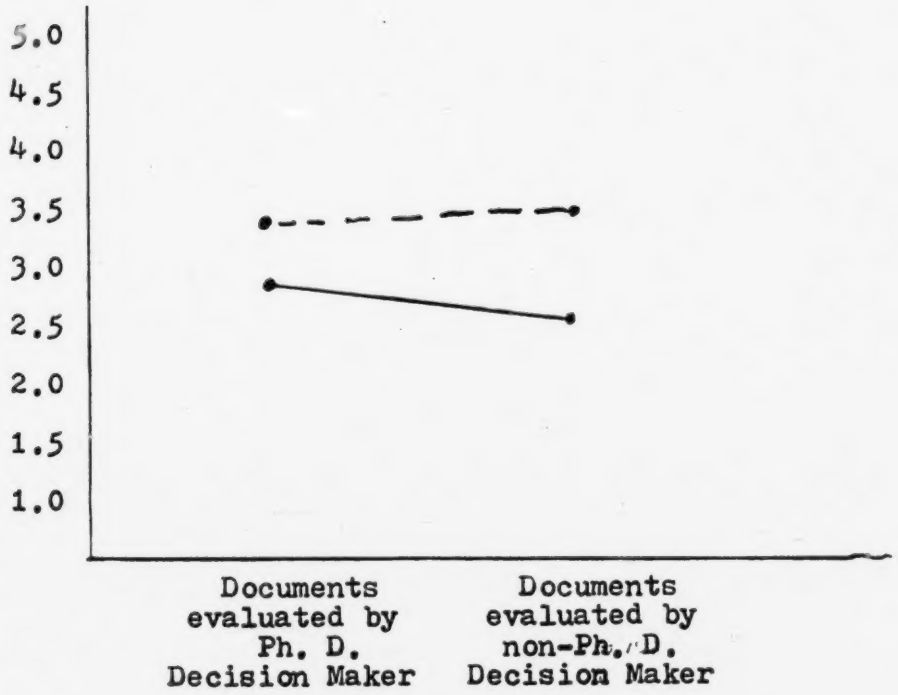
** Based on the research reports actually available. Reports not available to the reviewers were excluded from this computation.

*** Based on the total number of documents originally examined. Thus this figure indicates the percentage of the original documents which are supported by empirical data. The other reports either contained no data or were unavailable.

TABLE 14

REASONS FOR "REVISE" AND "REJECT" RATINGS

Reseach Specialists		Decision Makers	
Revise (n = 37)		Revise (n = 39)	
Unclear design	12	Weak or poor design	16
Representativeness of sample	12	Representativeness of sample	1
Too long	6	Too long	2
Redundant	1	Too complicated or technical	6
Incomplete analysis	3	Unclean design	5
		Incomplete analysis	2
		Insignificant to education	1
		Conclusions don't follow evidence	1
		"Do it over my way"	1
		No reason given	4
Reject (n = 52)		Reject (n = 23)	
Poor design	43	Poor design	12
Unclean design	4	Low generalizability	3
Insignificant to education	5	Vague or unclear	2
		Insignificant to education	6



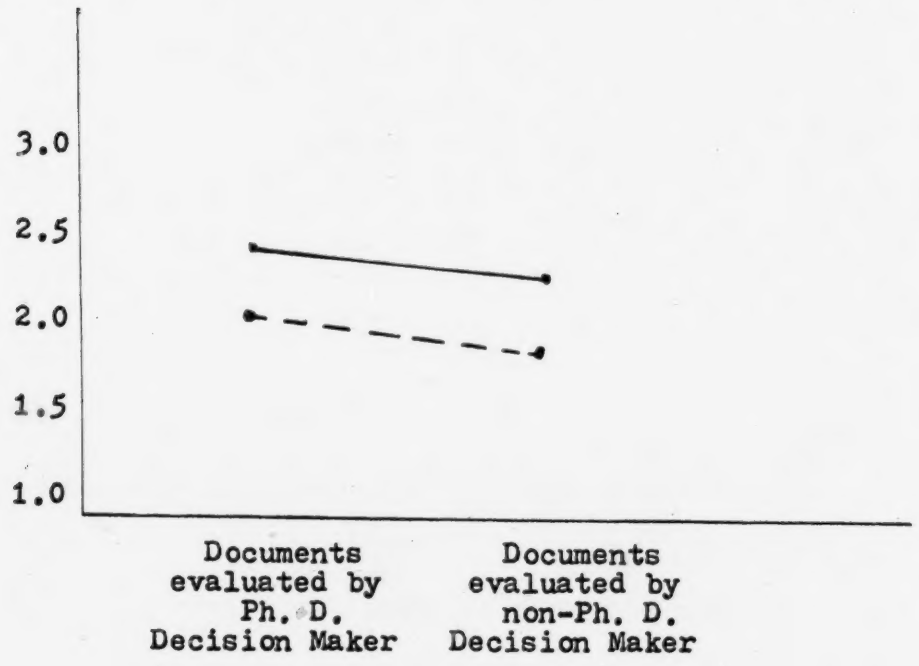
* Exact means for each group can be found in Table 8.

Also note that the directionality of Rating Scale A is reversed. A high rating on this scale indicates high quality in the document. On all other scales a higher rating indicates lower quality.

FIGURE 1

INTERACTION OF LEVEL OF EDUCATION OF DECISION MAKER

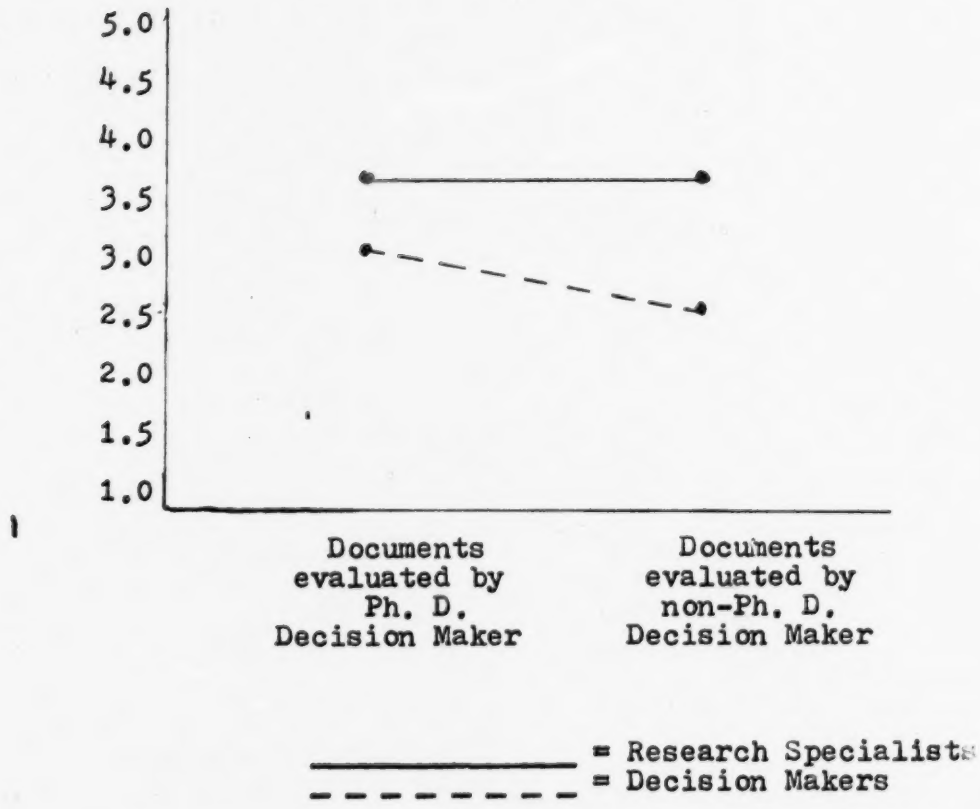
WITH DISCREPANCIES IN RATINGS ON RATING SCALE A *



* Exact means for each group can be found in Table 8.

FIGURE 2

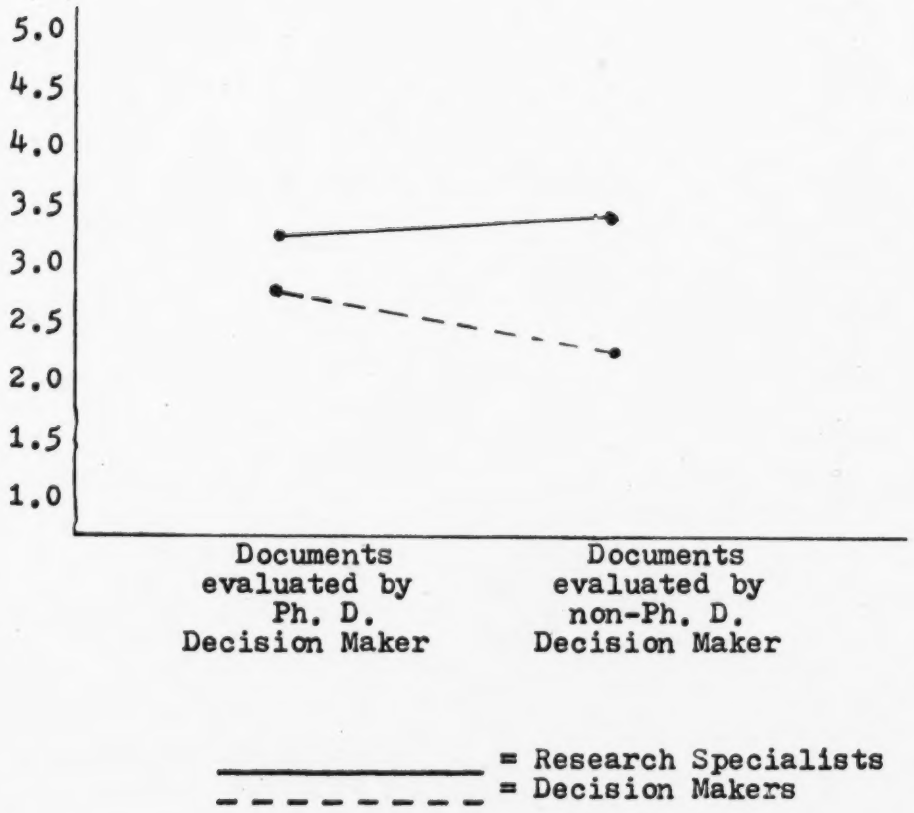
INTERACTION OF LEVEL OF EDUCATION OF DECISION MAKER WITH DISCREPANCIES IN RATINGS ON RATING SCALE B *



* Exact means for each group can be found in Table 8.

FIGURE 3

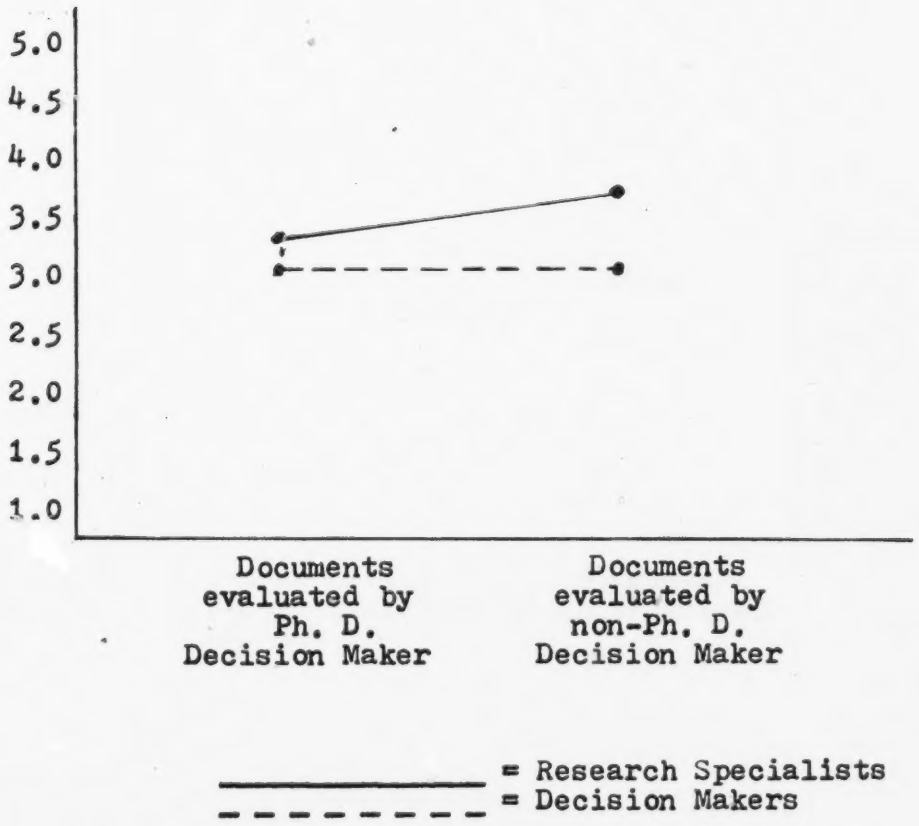
INTERACTION OF LEVEL OF EDUCATION OF DECISION MAKER WITH DISCREPANCIES IN RATINGS ON RATING SCALE C *



* Exact means for each group can be found in Table 8.

FIGURE 4

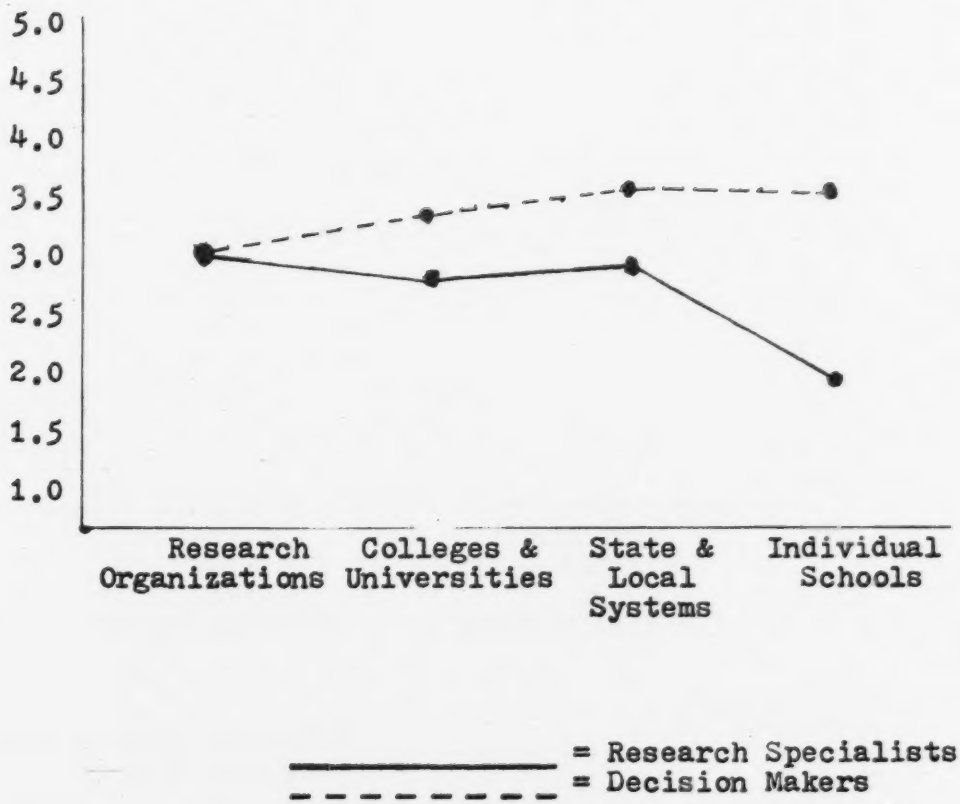
INTERACTION OF LEVEL OF EDUCATION OF DECISION MAKER WITH DISCREPANCIES IN RATINGS ON RATING SCALE D *



* Exact means for each group can be found in Table 8.

FIGURE 5

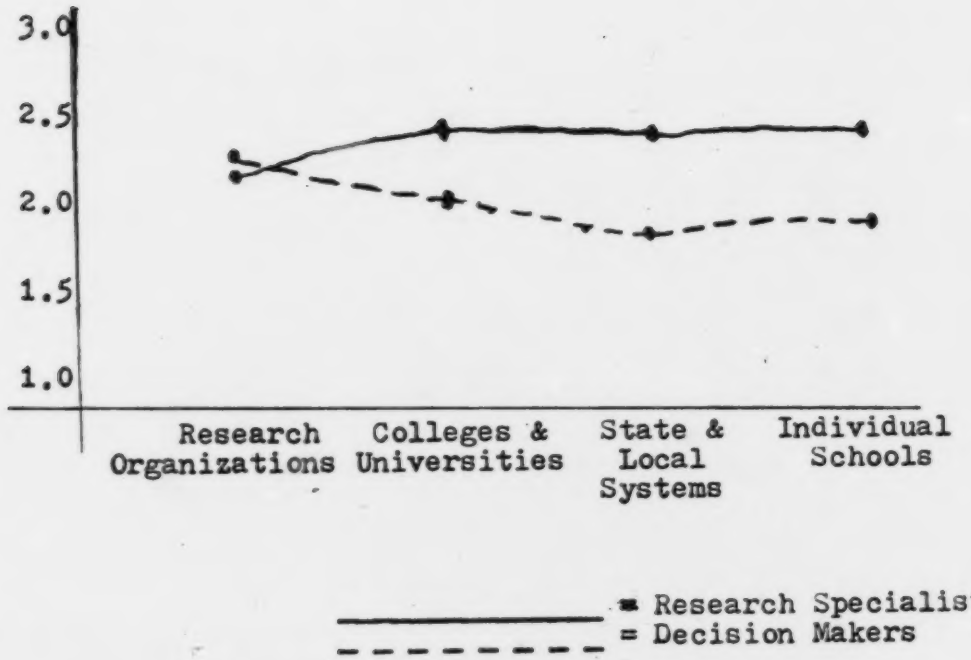
INTERACTION OF LEVEL OF EDUCATION OF DECISION MAKER WITH DISCREPANCIES IN RATINGS ON RATING SCALE E *



* Exact means for each group can be found in Table 9.

FIGURE 6

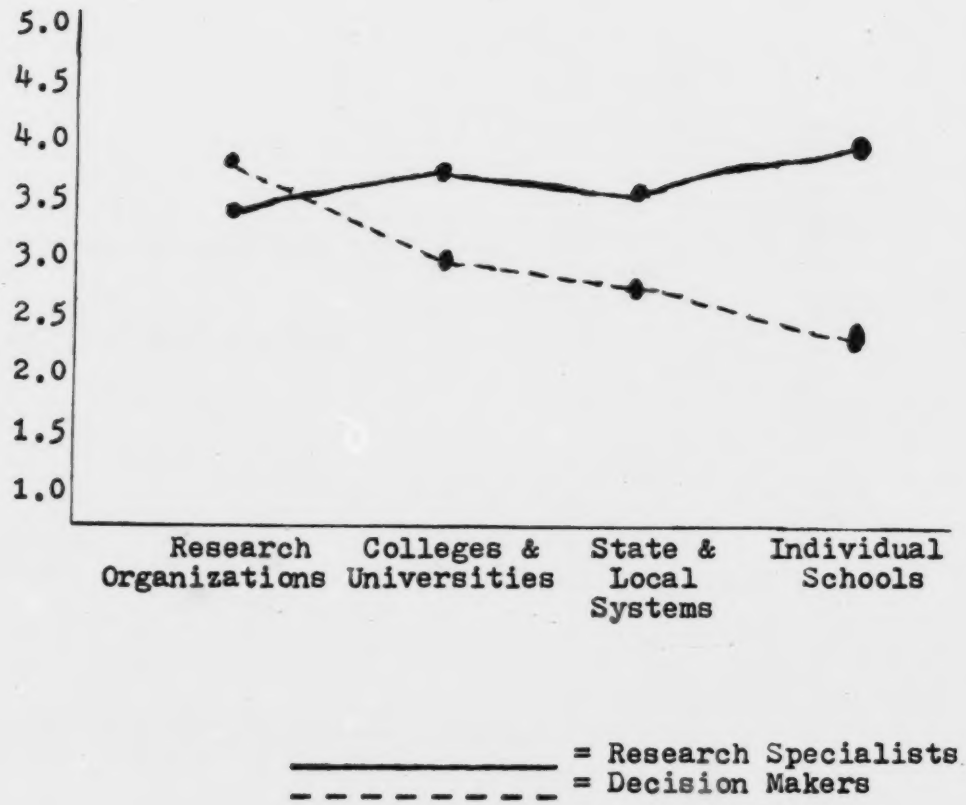
INTERACTION OF INSTITUTIONAL AFFILIATION OF DECISION MAKER WITH DISCREPANCIES IN RATINGS ON RATING SCALE A *



* Exact means for each group can be found in Table 9.

FIGURE 7

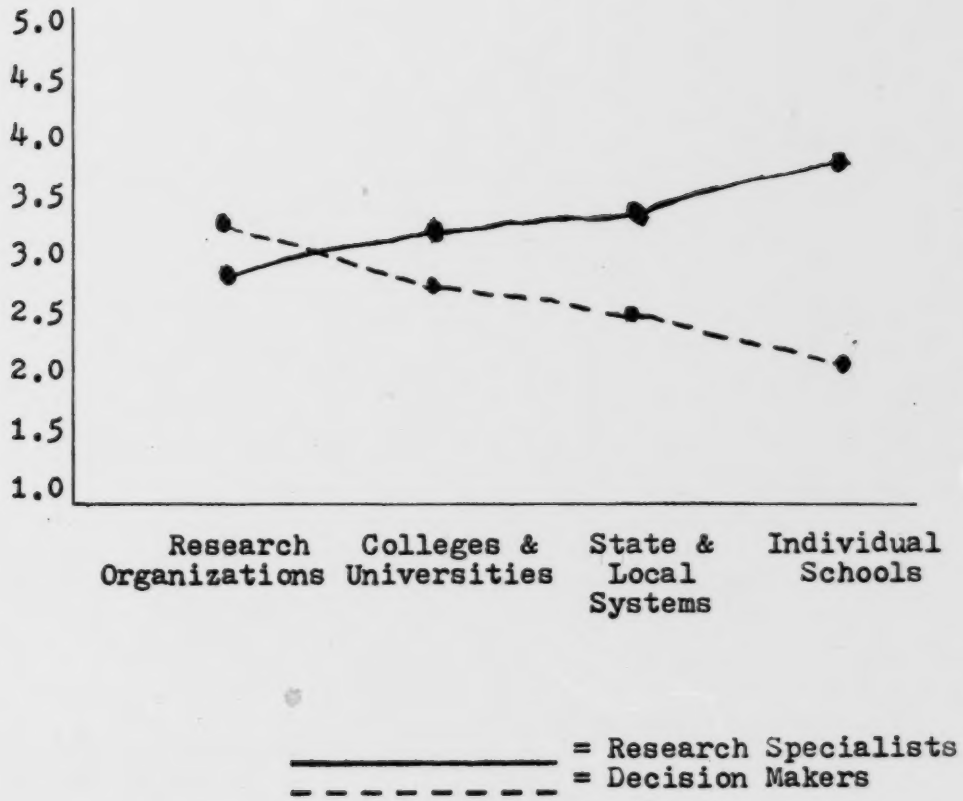
INTERACTION OF INSTITUTIONAL AFFILIATION OF DECISION MAKER WITH DISCREPANCIES IN RATINGS ON RATING SCALE B *



* Exact means for each group can be found in Table 9.

FIGURE 8

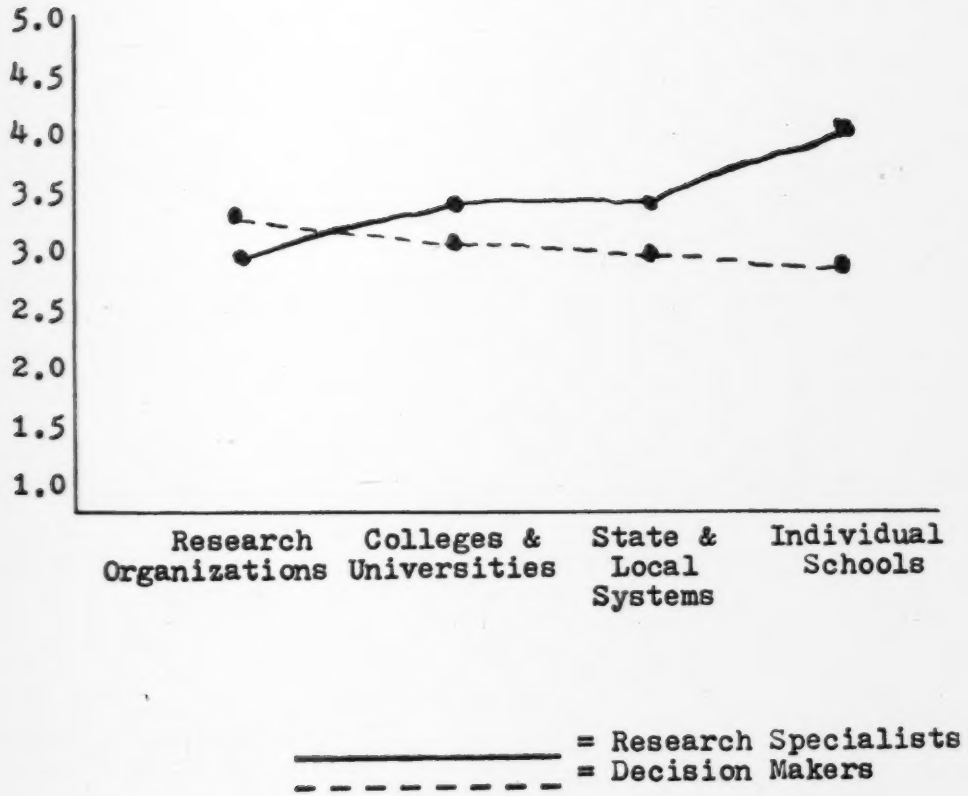
INTERACTION OF INSTITUTIONAL AFFILIATION OF DECISION MAKER WITH DISCREPANCIES IN RATINGS ON RATING SCALE C *



* Exact means for each group can be found in Table 9.

FIGURE 9

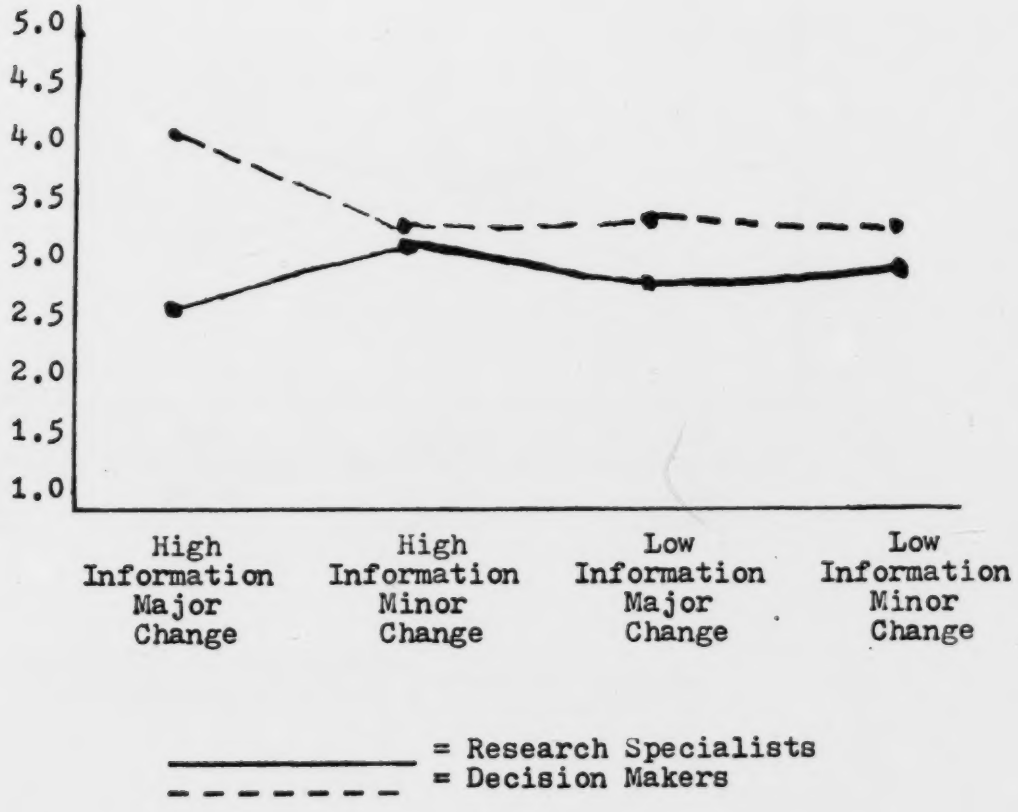
INTERACTION OF INSTITUTIONAL AFFILIATION OF DECISION MAKER WITH DISCREPANCIES IN RATINGS ON RATING SCALE D *



* Exact means for each group can be found in Table 9.

FIGURE 10

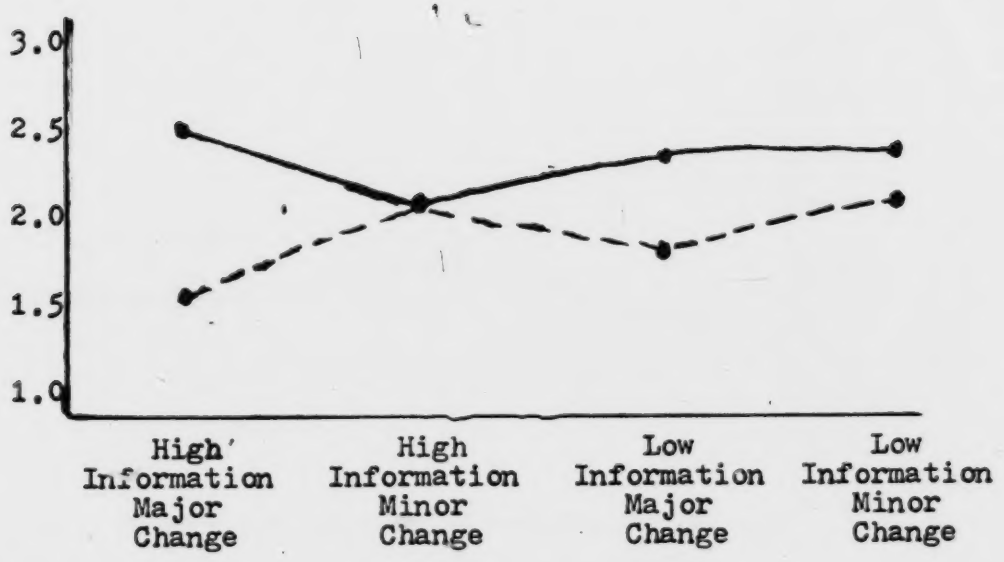
INTERACTION OF INSTITUTIONAL AFFILIATION OF DECISION MAKER WITH DISCREPANCIES IN RATINGS ON RATING SCALE E *



* Exact means for each group can be found in Table 12.

FIGURE 11

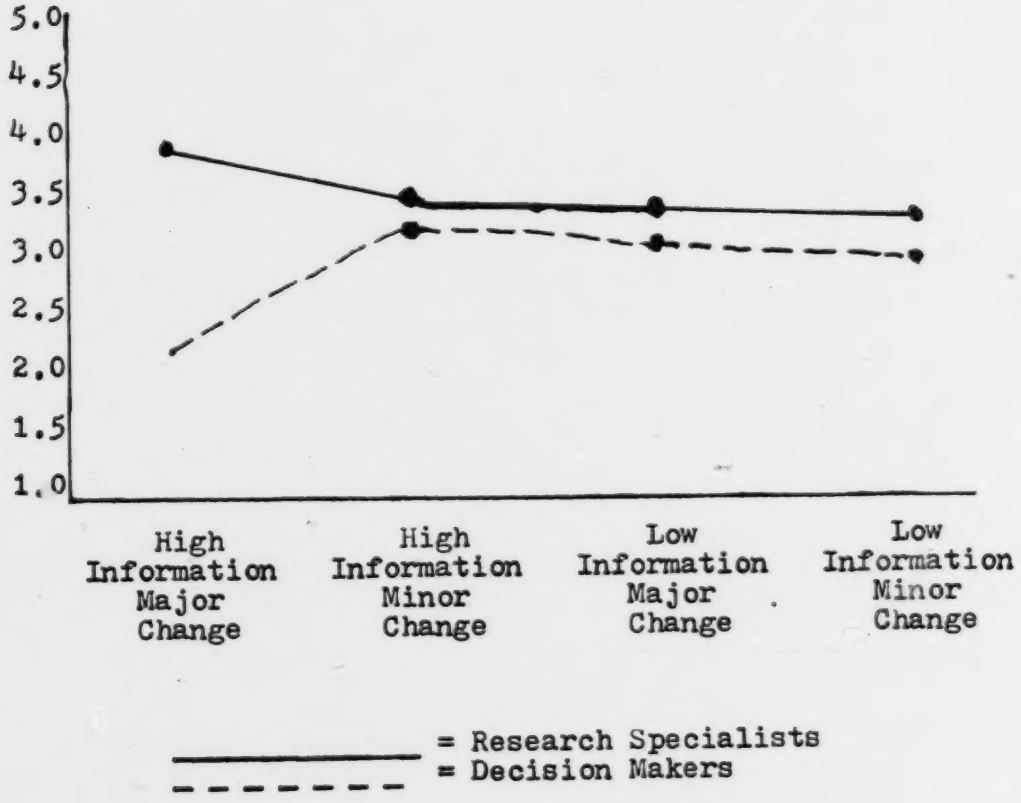
INTERACTION OF DECISION SETTING WITH DISCREPANCIES IN RATINGS ON RATING SCALE A *



* Exact means for each group can be found in Table 12.

FIGURE 12

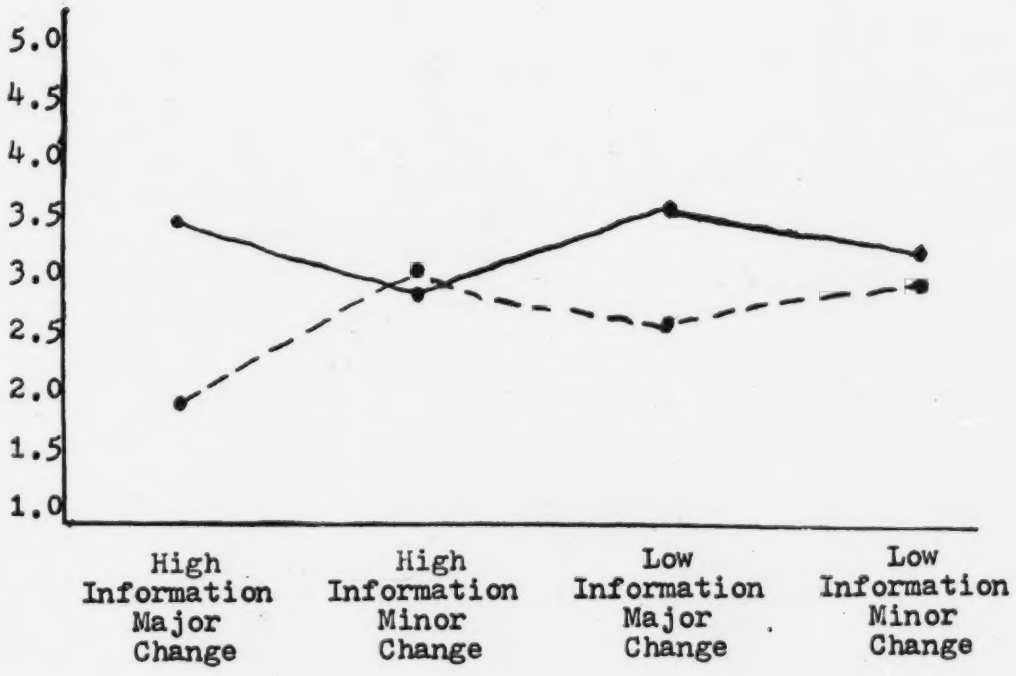
INTERACTION OF DECISION SETTING WITH DISCREPANCIES IN RATINGS ON RATING SCALE B *



* Exact means for each group can be found in Table 12.

FIGURE 13

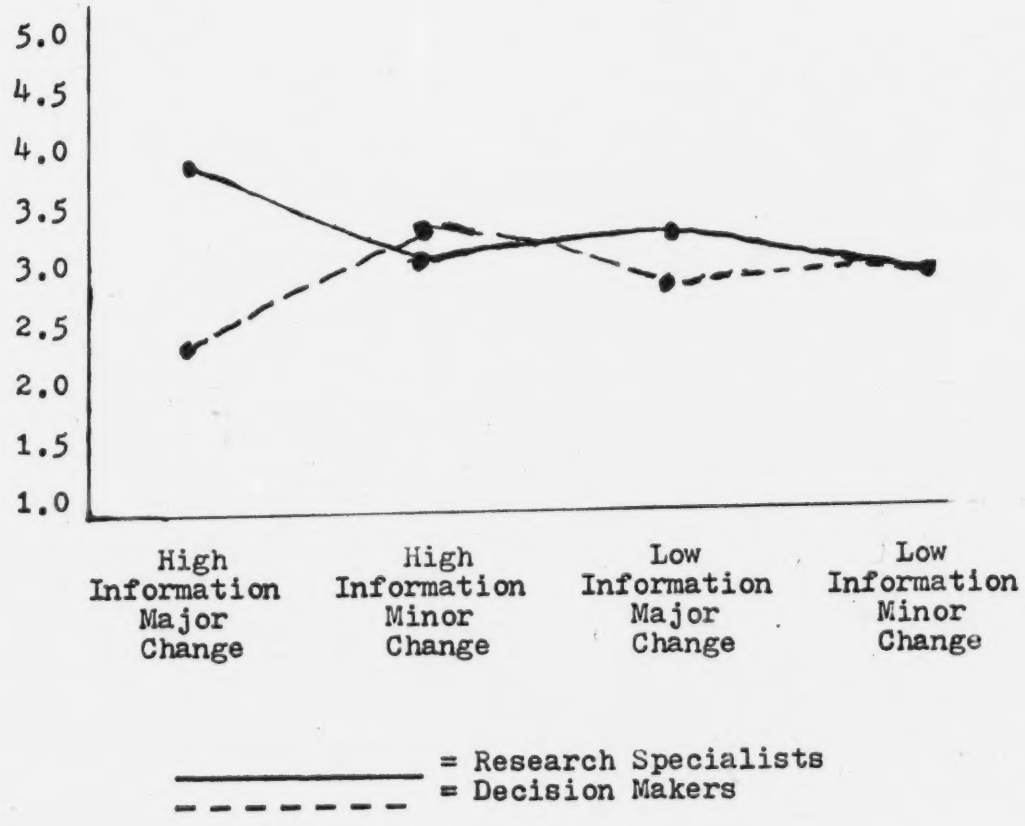
INTERACTION OF DECISION SETTING WITH DISCREPANCIES IN RATINGS ON RATING SCALE C *



* Exact means for each group can be found in Table 12.

FIGURE 14

INTERACTION OF DECISION SETTING WITH DISCREPANCIES IN RATINGS ON RATING SCALE D *



* Exact means for each group can be found in Table 12.

FIGURE 15

INTERACTION OF DECISION SETTING WITH DISCREPANCIES IN RATINGS ON RATING SCALE E *